

Leadership at Johannesburg

Political difficulties may stifle the impact of next week's sustainable-development summit in South Africa, but researchers and others must continue to pursue solutions to sustainability issues despite a lack of direction from governments.

Whether they are attending the gargantuan World Summit on Sustainable Development or watching from the sidelines, scientists committed to addressing the meeting's issues should not lose sight of their capacity to make a difference.

The world still has a long way to go in tackling the challenge of achieving economic development and improved quality of life while conserving resources and the environment. But to a degree, it was a tiny but tenacious minority of scientists who first defined the challenge and have done much to address it. Researchers put climate change and the Earth's finite resources on the political agenda, and their voices were some of the loudest in calling for policies to be changed in response to these threats. Since the Earth Summit in Rio de Janeiro in 1992, these researchers have been more closely in tune with both the public and policy-makers, and have begun to coordinate themselves to pursue sustainable development in many more arenas. Yet there remains much more that individual researchers and their institutions can do.

It is true that many of the obstacles to a sustainable future are political. The Bush administration's resistance to treaties such as the Convention on Biological Diversity and the Kyoto Protocol is a serious impediment to progress. And although climate and habitat loss are issues of critical importance, many others, such as energy supply and the security of food and water, also merit urgent attention. In obstructing a coordinated global response to these challenges, the US government is holding the sustainable-development agenda to ransom.

Other governments are far from blameless. The European Union (EU) is making the right noises about sustainable development — most notably by increasing funding for research in developing countries under its Sixth Framework Programme — but it could do far more to commit to the ideal. It could, for example, make more credible moves towards dismantling its Common Agricultural Policy, which undermines agriculture in developing countries by lavishing subsidies on European farmers. And in science funding, long-term support for the multidisciplinary research that is needed to address sustainable-development problems remains difficult to secure. Even the Millennium Ecosystem Assessment — a paragon of how research can be brought to bear on the issue — has secured only four years of funding.

Rise to the challenge

Given the scale of the task, the conflicting interests involved and the short-term nature of politics, next week's summit could well be a public flop. But researchers should not despair. To throw up their hands and blame the problem on 'politics' would be folly. With a little help, there is a lot that science can achieve.

Take the Montreal Protocol, signed in 1987, which banned ozone-depleting chemicals. The problem of ozone depletion was initially identified as such by the scientific and environmental community. High-profile meetings and action by non-governmental organizations then convinced the relevant industries of the problem, even before there was conclusive scientific evidence. Seeing which way the wind was blowing, the industries came on board, and to a large degree the decision to ban chlorofluorocarbons (CFCs) was already made by the time international policy-makers signed it into law.

There are parallels here with Kyoto — companies that use large

amounts of fossil fuel are clamouring for solid political commitments on carbon emissions. Industry is understandably reluctant to invest in infrastructure that may be illegal in a few years' time. Actions that change behaviour without firm commitments from government are known as 'type-2 measures' in the arcane parlance of the Johannesburg meeting, and have great potential to allow scientists and industry to address some of the critical issues of sustainable development while policy-makers are still plucking up the courage to legislate on them.

There is much that scientific expertise can achieve — especially in deploying existing know-how in the places where it is most needed. There is an abundance of high- and low-tech solutions to water management and energy generation in countries that lack advanced infrastructure, yet they need to be put in place in a rational way. Rich countries also have much to gain from sustainable-development research. For example, the development of coastal areas in the United States is regulated largely by local authorities, yet the impacts of such development are often felt great distances away. Fishermen and farmers can be introduced to scientifically informed approaches to fisheries management and agriculture where policies are lacking.

Researchers' role

On an individual or institutional level, researchers can begin to foster relationships with their colleagues in poor countries and to look for ways to apply their research to sustainable development. The idea is not that science should dictate policy — that is the job of governments. But science can act, and act effectively, outside the political arena.

Progress is already being made. Several multidisciplinary projects that are well suited to informing sustainable policy decisions have been created (see News Feature, page 812). Examples include the Millennium Ecosystem Assessment, projects led by scientists such as Pamela Matson in Mexico's Yaqui Valley, and research at the Khatmandu-based International Centre for Integrated Mountain Development — a partnership of scientists in eight Himalayan countries. There are commendable efforts in developed countries too, such as the research that contributed to Europe's Convention on Long-Range Transboundary Air Pollution, and the University of Southern California's PhD programme on sustainable cities, which focuses on Los Angeles.

Plans are also afoot to develop much-needed scientific knowledge and research capacity in the developing world. Efforts to introduce more fuel-efficient cooking stoves, funded by the World Bank, have begun to pay off by reducing biomass burning and respiratory disease in places such as China and India. The partnership between the University of California, Berkeley, and Nairobi-based Energy Alternatives AFRICA to establish a photovoltaic electricity industry in Kenya is now spilling over into other African countries. The EU-funded European and Developing Countries Clinical Trials Programme, which is expected to be unveiled at Johannesburg, could attract developed countries to carry out the high-quality clinical trials for locally important drugs that are so desperately needed in Africa.

Progress towards sustainable development will continue with or without effective guidance from next week's summit. Scientists, working in concert with others, are showing that they can help to steer the world towards a more sustainable future. ■





Tale of woe
NASA loses track
of its comet-seeking
craft
p806



Sense of scale
Initiative launched
to follow marine
migration
p807



Maelstrom
Experts clash over
evidence in US
anthrax case
p808



Cola wars
Rock paintings
spark outcry
from geologists
p809

Environmental impact tops list of fears about transgenic animals

Erika Check, Washington

A committee convened by the US National Academy of Sciences has expressed strong concerns that transgenic animals could damage their wild surroundings.

The panel was asked by the Food and Drug Administration (FDA) to list the top scientific concerns associated with animal biotechnology. In its 21 August report, the committee put environmental concerns about the impact of genetically modified animals on natural ecosystems at the top of the list.

One US company has already produced transgenic salmon that can reach adult size three times faster than other domesticated salmon. Some scientists and environmental groups have voiced fears over the fish, saying that they could outcompete wild salmon or introduce unwanted genes into native stocks. The FDA is reviewing the fish and has not yet said whether it will allow them to be sold by the company, Aqua Bounty Farms of Waltham, Massachusetts (see *Nature* **406**, 10–12; 2000).

But the company says that it is already gathering sets of data that will meet the



Transgenic salmon grow much faster than normal, but their impact on ecology remains unknown.

report's main concerns. "The only animal we intend to put on the market is a non-reproducing female salmon and that's intended to address an awful lot of these issues," says the firm's vice-president, Joseph McGonigle.

The academy's report says that there is not yet enough scientific evidence to prove that

transgenic salmon would be environmentally safe. There are also no definitive data on how safe the salmon — or other transgenic animals — are to eat, it notes. But the panel was less concerned about food safety as regulatory agencies already know how to remove toxic or allergenic products from the food supply.

In contrast, there is no good way to remove invasive species from their habitats if they cause problems, says panel chair John Vandenbergh, a zoologist at North Carolina State University in Raleigh. "I like to go fishing, but I don't know if I could catch all the genetically engineered salmon in the North Atlantic if they escape," he says.

In its report, the panel says that it has a "moderate level" of concern that proteins introduced into transgenic animals could cause allergic reactions if eaten by people. But it found no evidence that food products from cloned livestock — such as milk from cows — would prove dangerous.

Biotechnology companies that produce transgenic farm animals say that they hope the report will calm their investors. But environmental groups hailed the report as a critique of current US regulations on transgenics. "This backs up our contention that federal agencies aren't able to fully address the environmental and safety issues associated with these technologies," says Joe Mendelson of the Washington-based Center for Food Safety. ■

Public-access group plans journals

Kendall Powell, Washington

The Public Library of Science (PLS) — a group of researchers who last year threatened to boycott scientific publishers unless they put their journals online for free — will unveil its own publishing venture by the end of the year, one of its leading members says.

Michael Eisen, a geneticist at the University of California, Berkeley, and a founding member of the PLS, says that the venture will produce free-access print and online journals, covering costs by making page charges to authors.

A year ago, the PLS withdrew its plan to initiate a boycott of established journals from 1 September 2001 — despite obtaining few concessions from publishers.

Eisen says that PubMed Central, a free

archive established by the National Institutes of Health in 2000 for access to published biomedical research, is "woefully inadequate" in meeting researchers' needs.

Ed Sequeira, who manages PubMed Central, says that retrievals from the archive have doubled in the past year to 300,000 per month. He says that it will soon add more journals to the 80 currently archived. But the *Proceedings of the National Academy of Sciences* has revised the time lag from its original publication to free appearance on the archive from one month to six months.

And in a setback for the concept of publicly funded research archives, the US Department of Energy says that it is considering closing its PubScience search service for physical-sciences research. ■

Comet mission collapses as craft disappears

Tony Reichhardt, Washington

Comet research in the United States is facing an uncertain future after mission managers all but gave up hope of salvaging the CONTOUR (Comet Nucleus Tour) spacecraft. The craft went silent on 15

August as it fired an onboard rocket to leave Earth orbit.

The NASA-funded spacecraft was built and operated by the Johns Hopkins University Applied Physics Laboratory (APL) near Baltimore, Maryland. It was supposed to have visited two comets in four years, beginning with Comet Encke in 2003, to take close-up photographs of their icy nuclei (see *Nature* **417**, 889; 2002). Now the \$159-million mission seems to be a total loss.

For US comet investigators, CONTOUR's demise is the latest and most calamitous in a long line of setbacks. The United States was the only major spacefaring nation not to send a craft to Halley's Comet in 1986, and several NASA comet projects have been cancelled since, including the Comet Rendezvous Asteroid Flyby a decade ago, and a US lander for Europe's Rosetta mission, which is expected to rendezvous with a comet in 2011.

CONTOUR had operated flawlessly in the weeks after its 3 July launch, circling Earth as its orbit shifted gradually into the proper alignment to head towards Encke. This 'indirect launch mode' technique, pioneered on this mission, was designed partly to allow the use of a cheaper launch vehicle. Once the orbit was aligned, an onboard rocket motor fired to boost CONTOUR into interplanetary space.

Ground controllers had expected to establish radio contact with the spacecraft

shortly after the engine fired, but they never received a signal. Then, one day after CONTOUR vanished, the Spacewatch asteroid-tracking telescope in Arizona photographed two small objects streaking away from Earth at a distance of about 460,000 kilometres, in roughly the position CONTOUR would be if the rocket had fired. Mission managers took this as evidence that the spacecraft had broken into two pieces after the rocket fired, or that some part had come detached.

With no telemetry data or additional pictures, it will be difficult to determine what happened to CONTOUR. Although radio and optical telescopes continued searching the sky, APL managers were not optimistic about regaining contact as of early this week.

Mission director Robert Farquhar of APL was non-committal about whether his team might repropose the mission to NASA. CONTOUR was funded under the agency's Discovery programme for planetary exploration, and its price tag was only half the typical cost of such a mission. But other cometary spacecraft are being prepared for launch in the next few years, and it would take the CONTOUR team at least that long to ready a replacement.

For Farquhar, CONTOUR's loss was especially frustrating. "I wrote my first paper on getting to Encke's comet in 1972," he told the press last week. "I've been at this for 30 years, so I would like to have gotten some images." ■

Out with a bang: NASA's comet-seeking craft CONTOUR may have met a premature end.



Nobel laureate's dream for brain research finds home

Alison Abbott, Rome

An Italian competition to host a European Brain Research Institute (EBRI) has been won by a private hospital in Rome.

The institute is the brainchild of Nobel laureate Rita Levi-Montalcini, who personally organized the national competition without waiting for support or financial backing from the Italian government or any other institution.

The EBRI will be housed in a new building at Rome's Santa Lucia Hospital, which specializes in rehabilitating patients with neuromotor problems.

But as yet the new institute has no guaranteed funding to cover its running costs. Officials at Santa Lucia are relying on the hospital's growing status in Italian research, and on the reputation of Levi-Montalcini, to attract funds.

"We are confident that funding for some positions will emerge in next year's budget," says Luigi Amadio, director of the Santa Lucia Foundation, which owns the hospital. He also hopes to win support from the European Union.



A new building at Rome's Santa Lucia Hospital will house the European Brain Research Institute.

Levi-Montalcini, now 93, first mooted the concept of the EBRI last September, describing it as a truly European centre to attract foreign scientists into Italy, as well as offering Italian neuroscientists a chance to return home.

"I was surprised by the enthusiastic response from the scientific community," says Levi-Montalcini. Sites in Rome, Turin and Varese, near Milan, applied to host the centre.

The successful application was spearheaded by Amadio and neurologist Giorgio Bernardi of Rome's Tor Vergata University, who also works at the hospital. In recent years, Amadio and Bernardi have

greatly expanded research at Santa Lucia, and in 1992 it was awarded a 'research hospital' status, which allows it to apply for additional funds from the health ministry. Since then its research income has risen from 250,000 euros (US\$245,000) to 9 million euros last year, and it now employs 125 researchers.

Levi-Montalcini is also establishing an international scientific committee for the EBRI to recruit scientists, who will be able to establish research groups in any neuroscience field. "I'd like to see a holistic approach at the EBRI, with lots of cross-fertilization," she says. ■

Census of marine migration launched to gauge fish stocks

Rex Dalton, San Diego

Are there plenty more fish in the sea? The question continues to reverberate around most of the world's fishing grounds, but scientific data on the subject can sometimes prove elusive.

That's why researchers whose interests include species conservation and oceanography, as well as fisheries management, are keen to get more reliable information on the movements of marine animals. This month, on the Pacific seaboard of the United States, marine biologists took a big step towards their goal of understanding these migrations.

Researchers returned to San Diego from the ocean on 17 August after snagging and electronically tagging 80 tuna as part of the Tagging of Pacific Pelagics (TOPP) project, which aims to track bluefin tuna, several types of shark and albatross, elephant seals, leatherback sea turtles and other free-ranging species.

TOPP is part of an ambitious global undertaking called the Census of Marine Life, funded by the New York-based Alfred P. Sloan Foundation, the US Office of Naval Research and several other agencies. The census is intended to chart the diversity, abundance and distribution of marine organisms in the world's oceans over the next decade.

"The tagging will create a data set by observing animals that will present a better picture of the ocean," says Jesse Ausubel, an environmental scientist at Rockefeller University in New York who coordinates the Sloan Foundation's funding of TOPP, which involves 60 biologists, oceanographers and computer scientists.

The tagging project is the first of about a dozen similar schemes around the world that will be carried out as part of the marine census, and exploits the latest technological advances in tagging devices that can be attached to animals. Its tags store data on species location, water temperature and depth. The project uses both archival tags, which are recovered when species are caught or found dead, and transmitting ones — some as small as a deck of cards — that send information regularly for analysis via a French satellite instrument called ARGOS.

Barbara Block, a marine biologist at Stanford University in California who has pioneered tuna-tagging projects in the Atlantic Ocean (see *Science* **293**, 1310–1314; 2001), is leading TOPP in the Pacific. "The fishing was so good we ran out of electronic tags," she says of this month's voyage.



Size matters: measuring and tracking tuna helps researchers assess the status of fish populations.

TOPP's participants have also launched tagging operations to track salmon sharks off Alaska, blue sharks off Los Angeles, humpback whales off Oregon, and albatross near Hawaii and the Galapagos Islands.

George Boehlert, director of Oregon State University's marine laboratory in Newport, who studies the physical environments in which marine species live, says that data obtained from the tagging will help researchers to understand why animals venture to certain oceanic regions.

The information will complement that obtained from several other exploratory systems. For example, elephant seals, which plunge to depths of more than 600 metres several times an hour when at sea, will be monitored by the Argo float system that is being deployed to chart deep ocean currents and their impact on climate (see *Nature* **415**, 954–955; 2002).

Researchers hope that the tagging of tuna and other species will reveal migratory routes that will help biologists to understand Pacific fisheries. The project brings together groups of researchers from fields that do not traditionally interact. "Everyone is now on the same playing field," says Block.

► www.toppensus.org

Controversial animal feed builds concrete career in construction

David Adam, London

Walls have ears, and soon they could have the burnt remains of snouts, bones and skin as well. No longer able to feed meat-and-bone meal (MBM) to farm animals, British researchers are assessing its potential as a construction material.

The Meat and Livestock Commission has asked engineers at the Building Research Establishment (BRE) in Watford, near London, to investigate whether ash from incinerated MBM can be mixed into concrete and used to fill holes, lay roads and even build houses.

About 2.5 million tonnes of MBM are produced in Europe each year as a by-product of the meat trade. MBM was used in animal feed in Britain until 1996, when the practice was banned because the feed was thought to be playing a role in the spread of mad-cow disease.

Slaughterhouses are now left with mountains of animal leftovers that they currently incinerate and pay to dump as landfill.

The industry is keen to find cheaper ways to dispose of this unwanted legacy. One option is to use the MBM ash as an alternative 'aggregate' — material such as crushed stone and gravel that is used to make concrete. Ash from domestic incinerators is already used to make concrete in some European countries.

"The construction industry could basically have the stuff for free if it helped to offset disposal costs," says Martin Grantley-Smith, head of planning at the Meat and Livestock Commission in Milton Keynes.

Next month, a BRE team will start testing the mechanical strength of concrete made with the animal ash, and will see if any substances leach out. Ash has a different density and chemical composition depending on the burnt material.

"MBM ash contains a lot of calcium phosphate, not surprisingly because that's what bones are made of," says Rod Collins, who heads the BRE team. He adds that the MBM ash could possibly be used in construction, but acknowledges that people might not accept materials containing MBM in housing.

Collins is confident that the infectious agent that may cause mad-cow disease will not be a problem, however, as it should be destroyed in the incinerator. "You can never be absolutely certain," he admits. "But we're as sure as we can be that the material would be safe."

Anthrax case provokes doubt among experts

Jonathan Knight and Erika Check

In the FBI's search for whoever mailed anthrax to various US targets last autumn, both the hunter and the hunted are working hard to marshal facts in their favour. In so doing, both have released information of questionable scientific merit, experts say, perhaps shedding more heat than light on the continuing investigation.

Steven Hatfill, a former biodefence researcher at the US Army Medical Research Institute of Infectious Diseases at Fort Detrick in Frederick, Maryland, has been the subject of a barrage of media reports linking him to the case although the FBI says that he is not officially a suspect. *Newsweek* magazine reported on 12 August that bloodhounds had "jumped and barked, indicating they'd picked up the scent" when investigators led them to Hatfill's girlfriend's apartment and a restaurant where he had eaten the day before.

This information could only have come from law-enforcement sources. But bloodhound experts poured cold water on the idea that such a response would reliably reflect the dogs' recognition of a scent taken from packages that had each been handled by many different people almost a year ago.

Public plea

Hatfill, meanwhile, publicly pleaded his innocence by claiming, among other things, that as a virologist he wouldn't have known how to handle anthrax, which is a bacterium. Hatfill made the claim during a 15-minute public statement on 11 August in which he bitterly complained that FBI and media scrutiny had made a "wasteland" of his life.

"I have never, ever worked with anthrax in my life," Hatfill said. "It's a separate field from the research I was performing at Fort Detrick." Hatfill's civil lawyer, Victor Glasberg, then emphasized the point by suggesting that "not too many people doing the investigation understand the difference between a virus and a bacterium".



But microbiologists and bioweapons experts say that the distinction between the skill sets required to work with the two types of bioweapon agents is a difficult one to draw. Although the methods involved in growing viral and bacterial agents and using them in weapons differ in important ways, the overlap is considerable, they point out.

As obligate parasites, viruses can only be grown in host cells. Researchers who work with Ebola virus and monkeypox, as Hatfill did at Fort Detrick from 1997 to 1999, culture mammalian cells and infect them with the viral agent.

By contrast, the anthrax bacterium, *Bacillus anthracis*, is easier to grow than most viruses, requiring only nutritive broth or agar plates. The only technical part involves coaxing it to form spores, the dehydration-tolerant form used in last autumn's mailings. But this would pose little problem for a competent virologist, says Mark Wheelis, a microbiologist and arms-control expert at the University of California, Davis.

Other tricks are required to turn bioweapons agents into fine powders, or aerosols, that can be inhaled, and these vary from agent to agent. Viruses, for example, must be treated with chemicals to stop them falling apart when dried. But the steps to making an aerosol of any agent share several critical elements. Cultures are sprayed into fine particles and then dried, or first dried and then milled.

Emphasizing the differences between organisms merely sidesteps the more crucial question of access, says Al Zelicoff, a bioweapons expert at Sandia National Laboratories in New Mexico. "The real issue here is how someone got hold of a large amount of aerosolizable anthrax," he says.

On the scent

Criminologists, meanwhile, are scratching their heads over the bloodhound claims that appeared in the press. Several experts contacted by *Nature* say that the anthrax letters were unlikely to yield an ideal scent sample. Jerry Nichols, president of the Law Enforcement Bloodhound Association, says that scent from a letter could be used months later only if it had been lifted immediately and properly sealed. Nichols adds that he knows of cases in which courts "overturned or would not allow such evidence because there were too many questions of possible contamination". The FBI has declined to comment on the bloodhound incident or on how the scent was obtained.

Research has produced conflicting claims about bloodhounds' reliability. The most encouraging results suggest that they can correctly identify a suspect out of a line-up around 85% of the time. But one study found



Not in the know: Steven Hatfill says that he lacks the expertise to unleash an anthrax attack.

that when the person whose scent was taken is omitted from a line-up, dogs choose someone at random almost half the time anyway.

"It's not that dogs can't do a scent match," says Larry Myers, a veterinarian who researches canine scent tracking at Auburn University's College of Veterinary Medicine in Alabama. "But we don't know how well they do it. There simply haven't been enough studies."

Nichols and others say that dogs can also respond to unintentional signals from their handlers, who should therefore not be told about where the dogs are expected to react. But as the anthrax investigation drags on, the chances of a thorough double-blind study being done on this or any other technical issues raised seem to be remote. ■

Flood havoc at Czech institutes leaves research floundering

Prague Last week's floods have set back the Czech Republic's attempts to bring its science base up to the level of its neighbours, say researchers working there.

The water caused massive damage at Charles University in Prague, where physics and mathematics buildings were still under water as *Nature* went to press. Six research institutes of the Academy of Sciences were also damaged. The Institute of Physics was flooded by two metres of water, and a 750-million-euro (US\$730-million) electron microscope at the Institute of Inorganic Chemistry has probably been destroyed.

"In the last decade the government has invested heavily in science and education to try to reach West European standards," says Jiří Niederle, head of the academy's Council for International Affairs. He is now afraid that science will miss out on rebuilding funds in favour of urgent humanitarian priorities.

Geologists find cola rock stunt hard to swallow

New Delhi Soft-drink giants Coca-Cola and Pepsi have infuriated Indian geologists by



On the rocks: painted slogans will sabotage geological study for up to 20 years, say experts.

taking their advertising war to the foothills of the Himalayas. Adverts for the two drinks have been painted along the Manali–Rohtang pass in the state of Himachal Pradesh, a geologically important region. The posters have rendered the pass useless for research, say geologists.

"The rock painting is an irresponsible and inexcusable act," says A. K. Sinha, director of the Birbal Sahni Institute of Palaeobotany in Lucknow. "Rocks in the Rohtang can reveal a lot about tectonic activity in the Himalayas and the presence of minerals.

We can't even take our students there now."

Restoring the rocks to their original condition by scraping off the paint is a "virtual impossibility", says Ashok Sahni, a geologist at Punjab University in Chandigarh. Natural erosion will remove the adverts, but this will take 15 to 20 years.

The Indian Supreme Court issued notices to both companies on 14 August asking them to explain why they should not be penalized.

Prestigious prizes for young maths masters

Beijing The 2002 Fields Medals — regarded by many mathematicians as their field's Nobel Prizes — have been awarded to two young academics for work on uniting different areas of mathematics.

Laurent Lafforgue of the Institute of Higher Scientific Studies in Paris has been recognized for his contribution to advancing the Langlands Program, which proposes connections between analysis — the branch of mathematics that investigates functions — and number theory.

Vladimir Voevodsky, now at the Institute for Advanced Studies in Princeton, has been rewarded for developing 'cohomology' theories for algebraic varieties, the common solutions to sets of polynomial equations. He has thus provided new insights into

number theory and algebraic geometry.

The gold medals are due to be awarded this week at the 2002 International Congress of Mathematicians in Beijing. The prize is awarded every four years to mathematicians under the age of 40.

Astronomers bathe in the afterglow of γ -ray display

Washington Telescopes around the world were hurriedly repositioned last week after NASA's High Energy Transient Explorer 2 (HETE-2) satellite identified a γ -ray burst in the constellation of Sagittarius.

Observers studied the X-ray and optical afterglow of the 13 August event after its position was posted on the Internet. The origins of such bursts are unclear, but they are thought to be caused by events such as 'hypernovae' — massive stars collapsing to form black holes — or by the merging of two neutron stars or black holes.

The burst was welcome news for the NASA team. The first HETE satellite was lost in 1996 after it failed to enter the correct orbit. Technical and budgetary problems have prevented HETE-2, launched in 2000, from living up to expectations that it would rapidly relay the coordinates of 16–20 bursts a year. This is the first HETE-2 sighting that has been followed up so quickly and extensively.

The two hours it took for ground-based observers to locate the optical afterglow is still a far cry from the record 22 seconds it took other satellites to pinpoint a burst three years ago. NASA's Swift satellite, due to be launched in 2003, is expected to pinpoint 150 or more such bursts a year.

♦ <http://space.mit.edu/HETE>

Protest bomb found in xenotransplant lab

Rome A 15-kg bomb was found last week at an Italian institute that breeds transgenic pigs for research into xenotransplantation.

The home-made bomb, which had failed to detonate, was left in the toilets of the Institute for Experimental Zootechnics in Modena in northern Italy, which is funded by the Ministry of Agriculture. No group has claimed responsibility, but the attackers painted "Stop xenotransplantation" on the wall of the building. There was no warning and no precedent for the attack.

The institute's researchers say they are very shocked. "Our work is very well controlled and very open," says Maria Luisa Lavitrano, head of the xenotransplantation programme, who frequently appears on television to discuss the institute's research. The work is continuing despite the threat, says Lavitrano. The facility is now under police protection.

Britain probes outbreak of killer virus among seals

London An epidemic of phocine distemper, which wiped out half of northern Europe's seals in 1988, has led the British government to launch a research project on the virus.

The current outbreak is thought to have started in May on Anholt Island, off the east coast of Denmark. It has since hit the Dutch, Belgian, Norwegian and French coasts. Some 3,000 dead seals have been counted, though some may have died from other causes. The virus, transmitted by coughing, spreads quickly and attacks the immune system. Seals usually die from secondary infections such as pneumonia and respiratory problems. Few are immune and there is no known cure — the 1988 epidemic killed 18,000 seals.

The disease reached British shores this month, and the government announced on 13 August that it will fund a £250,000 (US\$380,000) nine-month research project. The work will include sampling dead seals to establish infection levels, as well as monitoring the progress of the disease.



In danger: seals face deadly virus.

ANDREW FORSYTH/RSPCA

Wanted: scientists for sustainability

Few observers expect much political progress at next week's summit on sustainable development. But it could mark the start of a transformation in the way scientists deal with sustainability issues. Tom Clarke reports.

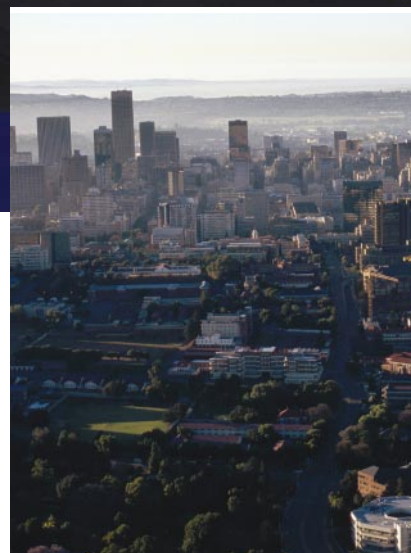
Pamela Matson (below, left) and Jane Lubchenco hope to bolster the role of sustainability research.



Politically speaking, the outlook for the World Summit on Sustainable Development is grim. Thousands of politicians, including the leaders of more than 100 countries, will fly into Johannesburg over the next two weeks — but most will have left their chequebooks at home. Delegates will spend ten days discussing how the world can continue to develop without jeopardizing the Earth's resources and life-support systems, yet even optimistic observers do not expect much in the way of new financial commitments.

But although the summit looks like being a political flop, it could be a turning point for scientists concerned with sustainable development. Whereas governmental attempts to address sustainability issues remain confused, researchers are slowly building a picture of what science can contribute. "The full power of what science has to offer has not even begun to be adequately tapped," says Jane Lubchenco, an environmental scientist and sustainability expert at Oregon State University in Corvallis.

At the summit, scientists like Lubchenco hope to hatch a new way of doing research that addresses the needs of sustainable development. How, for example, can fields from marine biology to economics work together to produce answers to specific local problems? And



Enlightening the world: researchers hope to make their presence felt in Johannesburg (above).

what can Western scientists do to help bolster research in developing countries, where local knowledge and scientific input into sustainability issues is often lacking? "Johannesburg is one of the first steps on an exciting and difficult journey," says Thomas Rosswall, executive director of the Paris-based International Council for Science (ICSU), which has been at the forefront of representing the interests of scientists at Johannesburg. "We need to create a science for sustainable development."

Ten years on

The conference was conceived as a follow-up to the last similar United Nations event — the Earth Summit, held in Rio de Janeiro in 1992. This time around, delegates will move beyond the environmental focus of Rio, and address how to protect human and natural resources while ensuring continued social improvement. Depletion of fresh water reserves, the use of unsustainable energy sources, food security, habitat loss

and global health — along with the important bearing that poverty has on these issues — are all high on the agenda.

The need to address these problems is clear. The six billion people currently living on Earth are thought to consume 40% of all terrestrial biomass, between a quarter and a third of marine resources and about 50% of the planet's accessible fresh water¹. In June, a report from the United Nations Environment Programme warned that half of the world's population is likely to suffer water shortages by 2032 (ref. 2). Yet the global population is projected to climb throughout this century, reaching at least 9 billion by 2100. "The trajectory we're on is not sustainable," says Lubchenco, who is currently ICSU's president-elect. "We are most definitely destroying the life systems of the planet."

Working out how to put the world on a sustainable track is far from clear. A sustainable approach to fisheries management may, for example, damage the food supply to local communities and effect international trade in fish. Fossil-fuel use might become more sustainable if, as some researchers suggest, the carbon dioxide produced is stored in the deep oceans — but is this simply passing the problem on to future generations? With so many different issues involved, there is rarely one definition of a sustainable solution. "Getting to grips with the concept is a bit like trying to catch soap in the bath," says John Lawton, chief executive of Britain's Natural Environment Research Council.

Science's summit

With political will in short supply at Johannesburg, government action is unlikely to clarify the situation. But many scientists are cautiously optimistic about the summit. The Rio meeting was organized by and for governments. Since then, growing networks of scientists, such as the Initiative on Science and Technology for Sustainability, based at Harvard University, have begun to address how to improve science's role in sustainable development.

Their efforts seem to be paying off, as scientists are now being invited to the party. Delegates at the summit's preparatory meetings consulted with researchers from a wide range of disciplines, and organizations such as ICSU have been asked to submit formal contributions to the summit. Together with the Third World Academy of Sciences in Trieste, Italy, and the World Federation of Engineering Organisations in Paris, ICSU will also be running a series of parallel meetings in Johannesburg. Early next year, the council hopes to summarize its findings in a report that could serve as a framework for science and sustainability. "We're no longer part of a fringe effort," says Rosswall.

High on the researchers' agenda is the problem of getting existing scientific expertise to where it is needed. The techniques,



Slippery issue: plans for sustainable fisheries management must also preserve local food supplies.

tools and experience to do the work already exist, argues Daniel Kammen, director of the Renewable and Appropriate Energy Laboratory at the University of California, Berkeley. Geographical information systems, which combine spatial information on a particular area with environmental, social and political data, have become more advanced and less expensive over the past decade. Global-positioning technology and satellite monitoring of the environment have also matured since the Rio summit.

The advances represent a unique opportunity to study climate, resource use and the relationship between health, the environment and development for the first time — but these tools are rarely focused on producing results that politicians can use. "A lot of this information never finds its way to policy-makers," says Felix Dodds, executive director of the Stakeholder Forum, a London-based group that supports non-governmental organizations and other parties working towards sustainable development.

Major international projects, for example, commonly have problems with creating such 'policy-relevant' science. The Intergovernmental Panel on Climate Change (IPCC) is widely admired for having communicated a strong scientific message amid the lobbying of industrial groups and environmental activists. But because the IPCC's reports are global in scope and contain huge amounts of data, it is difficult for policy-makers concerned with local issues to make use of them. "The crucial synthesis is linking local action to the global level," says Rosswall.

Local projects fare better, but almost all focus on single issues. The Consultative

Group on International Agricultural Research (CGIAR), which coordinates crop-research institutes worldwide, has set up regional crop-breeding and evaluation systems that are tailored to meet the needs of local farmers throughout the world. For example, the organization helped Cambodian farmers to rebuild rice production after Pol Pot was deposed in 1979, and to develop disease-resistant strains of important African and Asian staple crops such as bananas.

What is needed, say sustainability experts, is the expansion of local research projects so as to address the sustainability of all resources at a particular location at once. Such projects are beginning to emerge. Last year saw the launch of the Millennium Ecosystem Assessment (MEA) — an ambitious endeavour to assess the impact of factors such as shifts in land use and loss of biodiversity on the Earth's ecosystems (see *Nature* **417**, 112–113; 2002). Information on everything from fish stocks to nitrogen cycles will be produced, but the data generated are not just for ecologists. "The MEA focuses on things coming out of those ecosystems that people actually care about," says Lubchenco.

Local interest

An MEA study of Norwegian ecosystems should, for example, help the country's government to decide whether its fishing and oil-exploration industries can be expanded without damaging marine ecosystems. Other MEA members are providing technical support to studies in western China — an area earmarked for development, but which already suffers in parts from severe environmental problems such as soil erosion.

The project being run in Mexico's Yaqui Valley by Pamela Matson goes several steps further. By involving local researchers and politicians, she hopes to produce data that can directly influence the development of the region. During the 1990s, her team at Stanford University in California began



Thomas Rosswall says sustainability researchers face an "exciting journey".

studying the impact of increasing fertilizer use in the heavily irrigated valley, one of Mexico's most important agricultural centres. But run-off from upland agriculture feeds into the region, and water from the valley flows ultimately to the Sea of Cortez off the country's northwest coast, home to many birds, marine mammals and an important fishing industry. It soon became clear that studying the effect of fertilizer use in isolation from other factors would be of limited use.

Matson's project now involves local authorities, together with scientists and engineers from ten disciplines and five institutions in the United States and Mexico. The effect of land and chemical use, irrigation schemes and crop type on the local ecosystem are all being studied, as well as the impacts of external factors such as agricultural policies, globalized markets and drought. By combining this knowledge, researchers hope to reveal how these factors affect the terrestrial and aquatic environment and the income of farmers and city dwellers. "Decisions made in one part of the system, concerning one sector, cascade through the system, affecting many other sectors," says Matson.

The Yaqui Valley project is a work in progress, but Matson hopes that with more information — particularly on the human societies in the valley — a truly integrated approach should be possible. "If people there aim to make decisions that sustain their resources and their livelihoods, the place needs to be studied as an integrated system. We're trying to do that," she says.

Worldly wisdom

Projects such as Matson's could serve as a template for sustainability studies in other areas, and are certain to feature in the framework document being developed by ICSU and its partners. But implementing similar studies in other parts of the world could prove more difficult. Mexico has a good science base, but some countries in Africa, for example, would have difficulty in collaborating on a similar project. Discussions over how to improve scientific expertise and infrastructure in developing countries — 'capacity building' in the sustainability jargon — will be limited by the lack of new money available at the political summit, but the topic will be widely discussed at the parallel events.

Many sustainable-development organizations and governments are already aiding individuals and small projects in developing countries. The Swedish government, for example, is attempting to transfer its country's expertise in biomass fuels to the Baltic states. It hopes that use of the fuels, which are made from agricultural waste or specially planted crops, could reduce the Baltic region's reliance on coal and oil as energy sources.

Larger-scale projects are thin on the



Daniel Kammen, seen here with Samburu tribeswomen, says expertise must be sent to where it's needed.

ground, but one example is the European and Developing Countries Clinical Trials Programme, funded by the European Union. The 200-million-euro (US\$195-million) project, which is expected to be unveiled in Johannesburg, aims to establish centres for clinical trials at several locations in Africa by 2006. The scheme's backers hope that generating the right infrastructure and expertise will encourage pharmaceutical companies to run high-quality trials of treatments for diseases such as HIV, tuberculosis and malaria where they are most needed.

Marshalling the troops

Such projects are significant steps forward, but many researchers retain doubts about whether concepts such as policy-relevant science and capacity building can be successful on a large scale. As Lubchenco and colleagues put their framework document together, they face some major challenges. They will have to decide how individuals and institutes can best collaborate to study sustainability, for example, and whether sustainable-development science would benefit from coordination at a global level, perhaps by the United Nations. Any attempt to clarify these issues could be enormously helpful, especially to the many scientists who are interested in sustainability but are unsure how they can help. But even if ICSU and its partners can provide a blueprint for organizing the field, big political stumbling blocks remain.

Both politicians and science-funding agencies want to see a short-term return from most of their investments, and preferably in their own country or research area. Money for long-term multidisciplinary projects is hard to secure and often dries up after an initial outlay. Funding for the CGIAR centres from developed nations has, for example, been falling over the past decade. Other areas that are critical to sustainable development are also suffering. Research on renewable energy sources, for instance, has moved up the political agenda, but funding for the field has fallen by more than half as many developed countries slash their energy-research budgets³.

There are exceptions — funding agencies in some developed countries are beginning to make greater provision for research that focuses on sustainable development. But

overall the prospects are bleak. "A five-year electoral cycle is not equipped to deal with initiatives that will take 20 years to begin to take effect," says Lawton. And politicians are, by their very nature, unlikely to opt for large, long-term investments. "In a democracy it's hard to get turkeys to vote for Christmas," Lawton says.

Communication problems

But although individual researchers cannot overhaul political systems, they can ensure that politicians are aware of exactly how much science can contribute to sustainable development — which could in turn lead to changes in funding policy. "The limited actions towards sustainability that we already have are very largely driven by scientific data," says Jonathon Porritt, director of Forum for the Future, a London-based sustainability organization. "But science is not particularly good at selling itself." Communication about sustainability issues can be poor even among scientists, agrees Rosswall. "Many scientists can't properly explain what they are doing to other scientists," he says. "It's too easy to complain that it's the decision-makers who don't understand."

Avoiding such communication problems will be key to the success of the strategies being developed by ICSU and others. A solid description of exactly what researchers can do for sustainable development could provide a boost to everyone involved — from scientists in developed countries to administrators who hold the purse-strings of funding organizations. If those involved can make the right people realize what science can offer, they could help make to up for the likely political failings of Johannesburg. "It is time for us to break out of the ivory tower," says Rosswall. ■

Tom Clarke works in Nature's news syndication team.

1. Pimm, S. L. *The World According to Pimm: A Scientist Audits the Earth* (McGraw-Hill, New York, 2001).
2. UN Environment Programme. *Global Environment Outlook 3* (Earthscan, London, 2002).
3. Margolis, R. M. & Kammen, D. M. *Science* **285**, 690–692 (1999).

Web links

World Summit on Sustainable Development

♦ www.johannesburgsummit.org

International Council for Science

♦ www.icsu.org

Initiative on Science and Technology for Sustainability

♦ sustsci.harvard.edu/ists

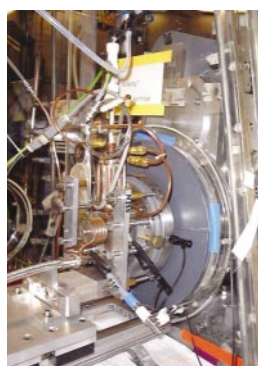
A very brief encounter

Little is known about the heavy elements that lie at the outer limits of the periodic table. But how do you investigate atoms that decay within seconds? Kendall Powell finds out. Kendall Powell finds out.

Christoph Düllmann cracked open a bottle of vodka when he detected his first hassium atom. Over the next few days, six more atoms registered on detectors at the GSI, a heavy-ion research centre in Darmstadt, Germany. Each one was a relief to Düllmann. He had spent two years developing the experimental game plan, and staked his doctoral thesis on the apparatus being able to study hassium.

Düllmann was not alone in feeling relieved. More than 30 researchers from 10 institutions had helped to plan how to probe the chemistry of hassium, a huge element that weighs in at number 108 in the periodic table and decays seconds after it is formed. They had gambled all on hassium conforming to the patterns of the periodic table — and the results, which appear on page 859 of this issue¹, show that they were right.

But researchers are finding that other heavy elements defy normal behaviour. The properties of most elements in the periodic table can be predicted by which column — or group — they appear in. But some heavy



Getting heavy: researchers at the GSI (above), and the equipment (inset) they used to detect the volatile tetroxide of the element hassium.

GSI

elements in the lower regions of the table buck this trend, throwing chemists' predictions into chaos.

Probing the chemistry of the heavy elements would be difficult even if they did play by the rules. Elements heavier than uranium have to be created artificially, usually by firing a beam of ions into a stationary target. Few collisions result in the creation of the desired element, so the process is very time-consuming. And the heavy elements tend to be unstable — many of them decay into lighter elements in seconds. So although physicists have extended the periodic table by a few new elements every decade since 1940, mapping out the territory as far as element 112, only in the past two decades have advances in target design and beam intensity allowed researchers to study heavy-element chemistry.

The quirky behaviour displayed by heavy elements has its roots in their electronic structure. In all large atoms, electrons orbiting close to the nucleus act as a shield between the attractive forces of the nucleus and the electrons in the outer orbits. But in the heavy elements, the high positive charge on the nucleus causes the electrons in the inner orbits to move at speeds close to that of light which, according to the special theory of relativity, increases their mass. This sends the electrons into even shallower orbits, increasing the shielding effect for those in the outer orbits.

The heightened shielding can change the way in which the outer electrons interact with other atoms and molecules, potentially

putting the element out of step with the other members of its chemical group. "One

would expect a major break from periodic-table trends," says Pekka Pyykkö, a theoretical chemist at the University of Helsinki. "But no one knows when it will happen. In what column, if any, will you see major changes?"

Risky business

These uncertainties are a major problem for nuclear chemists, who risk spending years designing equipment that, if an element behaves in an unexpected way, may not be able to detect what they were seeking. To succeed, sometime fierce rivals have had to collaborate, and everyone involved has had to rely on a little luck.

In the case of hassium, researchers had to wait 12 years before they could even begin work. The element was discovered in 1984 at the GSI, and takes its name from Hesse, the state in which the institute is based. Researchers there fired a beam of iron ions into a lead target. For a few of these collisions, an iron nucleus overcame the repulsion from a lead nucleus and the two fused to create hassium. But with a half-life of 1.5 milliseconds, the hassium isotopes were too short-lived to be used in experiments.

The breakthrough came in 1996, when GSI researchers were searching for element

GSI



Christoph Düllmann gambled that hassium would behave in line with the rest of its group.

Group theory: the properties of some of the heavier elements deviate from what would be expected from their position in the periodic table.

H 1	He 2
Li 3	Be 4
Na 11	Mg 12
K 19	Ca 20
Rb 37	Sr 38
Cs 55	Ba 56
Fr 87	Ra 88

Sc 21	Ti 22	V 23	Cr 24	Mn 25	Fe 26	Co 27	Ni 28	Cu 29	Zn 30	Ga 31	Ge 32	As 33	Se 34	Br 35	Kr 36
Y 39	Zr 40	Nb 41	Mo 42	Tc 43	Ru 44	Rh 45	Pd 46	Ag 47	Cd 48	In 49	Sn 50	Sb 51	Te 52	I 53	Xe 54
Lu 71	Hf 72	Ta 73	W 74	Re 75	Os 76	Ir 77	Pt 78	Au 79	Hg 80	Tl 81	Pb 82	Bi 83	Po 84	At 85	Rn 86
Lr 103	Rf 104	Db 105	Sg 106	Bh 107	Hs 108	Mt 109	H 110	Uub 112							

La 57	Ce 58	Pr 59	Nd 60	Pm 61	Sm 62	Eu 63	Gd 64	Tb 65	Dy 66	Ho 67	Er 68	Tm 69	Yb 70
Ac 89	Th 90	Pa 91	U 92	Np 93	Pu 94	Am 95	Cm 96	Bk 97	Cf 98	Es 99	Fm 100	Md 101	No 102

112. By firing a beam of zinc ions into a lead target, they became one of the first groups to succeed at creating the element, which currently has the temporary name ununbium. Amid the debris of the particles and nuclei that 112 decayed into, a hassium isotope with a half-life of nine seconds was discovered.

This started the ball rolling on the hassium chemistry study. Düllmann and his collaborators focused on the volatility of hassium tetroxide, the gas formed when hassium reacts with oxygen. Elements lower in the group form more volatile tetroxide gases, so hassium tetroxide was expected to be at least as volatile as osmium tetroxide, the gas formed by the element just above it in the group. This was backed up by work from GSI theoretician Valeria Pershina², who suggested that hassium tetroxide would follow the pattern of the other elements in its group, such as osmium and ruthenium.

In character

Magnesium ions were pounded into a curium target to create atoms of the long-lived hassium isotope, which were then flushed into a thin tube by a stream of helium and oxygen. The mixture was heated to 600 °C, which, if the predictions were right, should have prompted the formation of hassium tetroxide. The gas was then sent over a series of silicon detectors held at different temperatures. Tetroxide gases adsorb on silicon, and the temperature at which they do so can be used to infer the gases' volatility.

Crunch time came last May, as experiments were run to reveal whether the researchers had gambled their time and money on the right kind of experiment. Ten days into the study, the first hassium tetroxide molecule was detected — and it was within the expected temperature range. Düllmann double-checked the equipment and toasted others in the control room with a small drink of vodka. Hassium, it seems, does conform to the behaviour expected of it as a member of its chemical group.

But other heavy elements are less predictable. During the 1980s, experiments on dubnium, element 105, showed that it behaves similarly to niobium, the element two places above it in its group, rather than its nearest upstairs neighbour tantalum³. Rutherfordium (104) is also known to break patterns within its group^{4,5}, but others including seaborgium (106) follow some normal trends⁶.

Deviant behaviour

One element, however, looks set to throw the rulebook away completely. A team led by Alexander Yakushev at the Flerov Laboratory of Nuclear Reactions in Dubna, Russia, is currently trying to capture element 112. This element is in group 12, so it ought to behave like a more volatile version of mercury. But relativistic effects are expected to make some of its properties more like those of an inert gas such as radon⁷.

Yakushev and his colleagues decided to hedge their bets by building an experimental set-up to detect both mercury-like and radon-like activity. Unlike the hassium experiment, the group is assessing the chemistry of element 112 in its unbonded state. Calcium ions are smashed into a uranium target, and the atoms of element 112 that are created are flushed onto two different detectors. If the element behaves like mercury, it should bind to the gold-covered surface of one of the detectors. If it doesn't, it will be



All together now: the international team celebrates the end of the hassium experiment.

swept into an ionization chamber where its radioactive decay chain will be detected.

In preliminary, unpublished results, nothing stuck to the gold surface, but eight atoms were detected in the chamber, indicating that 112 may display radon-like properties. Other elements disrupt the trends within their groups, but 112 appears to be acting as if it belongs to another group altogether.

Like other work on the chemistry of heavy elements, the Flerov team relies heavily on techniques and predictions developed by researchers around the world. The hassium study was an equally multinational affair. Teams at Lawrence Berkeley National Laboratory in Berkeley, California, and the Paul Scherrer Institute in Villigen, Switzerland — where Düllmann is based — designed the detectors. Flerov researchers performed preliminary studies of osmium and various German groups made the target, organized the final experiment and provided theoretical calculations. "In principle, it was everyone working in gas-phase chemistry of super-heavy elements in the world," says Yakushev, who is a co-author on the hassium paper.

Such collaborative spirit does not always come easily. The competition between groups to find new elements is intense, partly because the discoverers get to name the element. But the complexity of the chemistry experiments forces the groups to work together. "It requires walking a fine line and tremendous diplomacy," says Heino Nitsche, a nuclear chemist at Lawrence Berkeley.

This delicate balance was damaged earlier this year, when Victor Ninov, a physicist on the Lawrence Berkeley team, was accused of falsifying some of the data behind his lab's claim to have detected elements 116 and 118. Ninov, who denies any wrongdoing, was sacked from his lab and the paper on the results was retracted⁸. Earlier studies in which he was involved at the GSI are also under investigation.

Yakushev calls the incident "a lesson for all of us", and says future results need to be available for others working in the field to check. With funding for basic research shrinking and use of particle beams reserved for high-energy physics, nuclear chemists worry that the scandal will hinder their efforts. With few immediate applications, their field can do without bad publicity, even if they are producing fascinating science. "The international group is too small to be pulling in different directions," says Nitsche. "The field cannot afford it when the fundamentals of the periodic table are in question."

Kendall Powell is an intern with Nature.

1. Düllmann, Ch. E. *et al.* *Nature* **418**, 859–862 (2002).
2. Pershina, V., Bastug, T., Fricke, B. & Varga, S. *J. Chem. Phys.* **115**, 792–799 (2001).
3. Gregorich, K. E. *et al.* *Radiochim. Acta* **43**, 223–231 (1988).
4. Czerwinski, K. R. *et al.* *Radiochim. Acta* **64**, 23–28 (1994).
5. Czerwinski, K. R. *et al.* *Radiochim. Acta* **64**, 29–35 (1994).
6. Schädel, M. *et al.* *Nature* **388**, 55–57 (1997).
7. Pitzer, K. S. *J. Chem. Phys.* **63**, 1032 (1975).
8. Ninov, V. *et al.* *Phys. Rev. Lett.* **89**, 039901(E) (2002).

Summit: vague answers to well-known problems?

Multinational negotiations can work, but not where local people are causing the problem.

Sir — Two types of environmental problem resist solution by multinational summits such as the Johannesburg Summit on Sustainable Development (see V. Gewin, *Nature* **417**, 475; 2002). First are local problems that cannot be managed through global mandates; second are global problems that impinge on domestic authority.

Many 'global' problems are actually local ones with global implications, such as deforestation and desertification. Multilateral negotiations are ineffective in this case: the Convention on Deforestation will not stop a poor farmer cutting down a tree for firewood; declarations on sustainable development will not change North America's appetite for consumption.

Second, summits sponsored by bodies such as the United Nations (UN) have a mandate to respect national sovereignty. But in many fields this conflicts with policy-making. Climate-change negotiations have been challenged by nations worried about

the effects on domestic policy. In multilateral trade agreements, the World Trade Organization (WTO) forbids differentiation of traded goods on the grounds of their production methods, so environmentally beneficial standards cannot become the basis of trade policy.

To address local problems: transferring technology and reducing poverty may be the best way of promoting good environmental management in poorer countries.

To address state sovereignty: the International Commission on Intervention and State Sovereignty's report, *The Responsibility to Protect* (www.iciss-ciise.gc.ca/report-e.asp), lays out the principles and conditions whereby this is subordinate to humanitarian causes. We can conclude that states have the same responsibility to protect their people from environmental threats as from military threats. If environmental problems affect many people, the international community should have the

mandate to mediate solutions. This could have helped the 50,000 in Malaysia affected by Indonesia's forest fires in the 1990s.

We also need strong national legislation to ensure that the environmental costs of business are reflected in prices, and that trade policy supports production methods that absorb environmental costs.

Judging from the preliminary draft documents for Johannesburg, all we can expect is another declaration about the problems we face, coupled with grand, imprecise solutions. The problems were well articulated 30 years ago at the first UN environment summit; surely it is time to assess the strengths and weaknesses of summits in areas where diplomacy fails to promote environmental governance and devise appropriate new strategies.

Evan D. G. Fraser, Warren Mabee

*The Liu Centre for the Study of Global Issues,
6476 Northwest Marine Drive, Vancouver,
British Columbia V6T 1Z2, Canada*

Physicists take issue of misconduct seriously

Sir — The News Feature "Time to wise up?" (*Nature* **418**, 120–121; 2002) concludes, in the light of recent allegations about results from Bell Labs, that physical scientists' organizations are not tackling misconduct.

You quote an excerpt from a statement by the American Physical Society (APS) and comments by an APS spokesperson. These do not provide an accurate picture of APS policy. Our 1987 statement (see www.aps.org/statements/87.1.html) comments on why physicists may not have felt a need for a misconduct policy in the past. But you do not mention our 1991 guidelines for professional conduct (see www.aps.org/statements/91.8.html), an omission that may give readers the erroneous impression that the APS has never adopted a code of ethics.

The 1991 guidelines cover the types of alleged misconduct now being investigated. They state: "Fabrication of data or selective reporting of data with the intent to mislead or deceive is an egregious departure from the expected norms of scientific conduct." The spokesperson's comment that the APS "has no specific plans to revise its misconduct policies at this time, but is always alert to making changes in the future" should be read in this context. If events show that these guidelines are incomplete, the APS is prepared to revise

its policies or take other appropriate action.

William F. Brinkman

*American Physical Society, 1 Physics Ellipse,
College Park, Maryland 20740, USA*

Fraud: the system works

Sir — Your News Feature (*Nature* **418**, 120–121; 2002) seems to try to shift the blame for the "Schön affair" to the scientific community. We have followed this work since the first publications. The answer to your question of whether we are ready to police ourselves is simple. The scientific system is working. Physical scientists do police themselves: not only is the scientific community ready to "tackle the issue of misconduct", it has already done so.

Problems with the Bell Labs data (see *Nature* **417**, 789; 2002) were uncovered by physical scientists asking tough questions months before journalists got a hint of them. The impression left by your article, that we were taken for a ride, could not be farther from the truth. Journalists have blown things out of proportion, while scientists were trying to reproduce results.

It is not easy to identify such problems and it takes time for the system to kick in. The results Schön *et al.* presented were not unreasonable: several groups have observed similar (although much smaller) effects over more than 20 years. It is still not clear whether there has been any fraud; the possibility exists, however unlikely, of an

experimental mistake. This will eventually be settled by the investigating committee.

Natural science has a self-regulating system that has evolved over centuries. Sooner or later the true facts about claimed physics phenomena are revealed by a worldwide iteration process by diverse research teams. This system works.

Ivan K. Schuller*, Gernot Guentherodt†

**Department of Physics, University of California,
San Diego, La Jolla, California 92093, USA*

*†II Physikalisches Institut, Rheinisch-Westfälische
Technische Hochschule, 52056 Aachen, Germany*

Impartial review is key

Sir — Your News Feature (*Nature* **418**, 120–121; 2002) raises important issues that the physical-science community must face. Argonne National Laboratory's code of ethics calls for a response very similar to that of Bell Labs, namely: "The Laboratory director may appoint an ad-hoc scientific review committee to investigate internal or external charges of scientific misconduct, fraud, falsification of data, misinterpretation of data, or other activities involving scientific or technical matters." Such a response, giving ample opportunity for all sides to be heard by a panel of disinterested experts, is a natural and effective means of dealing with cases of possible misconduct.

George Crabtree

*Materials Science Division, Argonne National
Laboratory, Argonne, Illinois 60439, USA*

The price of consumerism

Tackling the problems of consumption may require radical solutions.

Confronting Consumption

edited by Thomas Princen, Michael

Maniates & Ken Conca

MIT Press: 2002. 386 pp. \$67, £45.95 (hbk);

\$26.95, £18.95 (pbk)

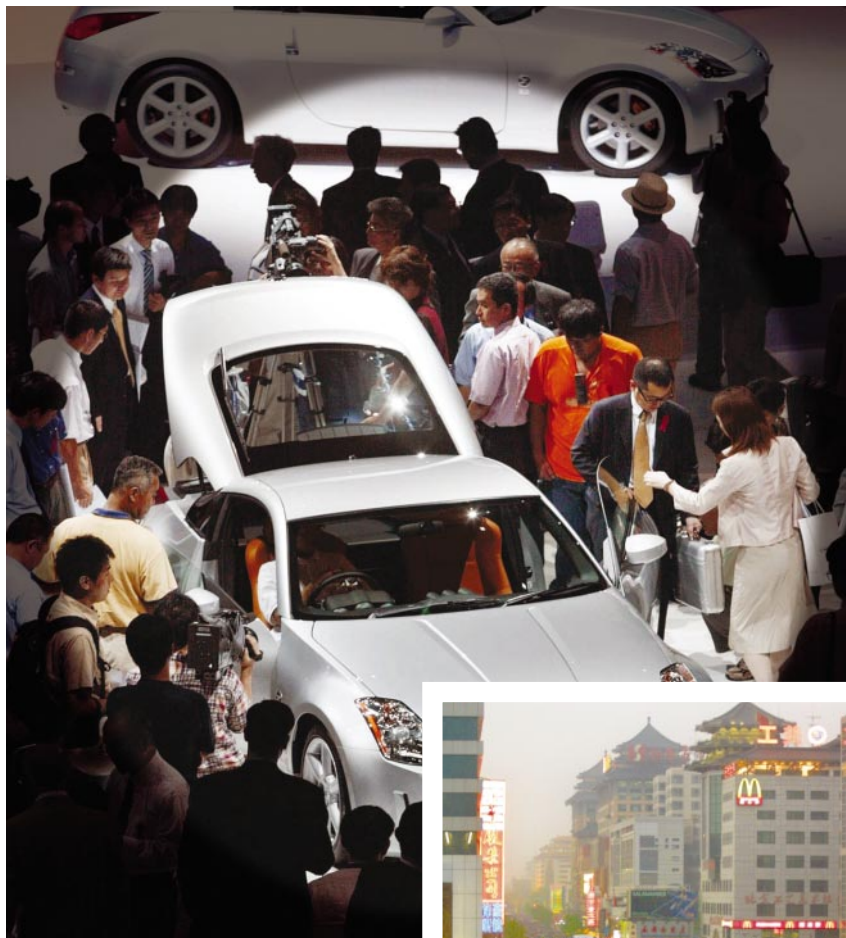
Norman Myers

An important new issue is emerging in the arena of sustainable development that rivals environmental concerns in importance. It is consumption — or rather, mis- and over-consumption. This is a less tractable problem than environmental degradation in that politicians see efforts to contain consumption as a superb way to lose votes. It is at the heart of the sustainability debate, as it covers the depletion of natural capital, such as key non-renewable resources, and the capacity to absorb pollution. But it will receive almost no attention at the World Summit on Sustainable Development in Johannesburg next week.

Nor, for that matter, has consumption received the attention it deserves from researchers in either the natural or the social sciences. Indeed, certain economists view consumption as the natural purpose of the economy. In sociological terms, we have only hazy ideas of why people consume as they do even when it ostensibly does not deliver the satisfaction it promises. Do consumers feel that they enhance their identity by what they purchase and by keeping ahead of the Joneses? Does shopping truly provide retail therapy? These questions remain largely a black hole of research, but we have even less idea of how to supply consumerist needs through other means.

Moreover, consumption is not an issue for rich countries alone. In developing and transition countries there are well over one billion people with enough income to enjoy an affluent lifestyle. Their aggregate spending in purchasing-power parity (as measured in local terms) already matches that of the United States. Certain effects of their consumerism, such as pollutant emissions from cars, which cause urban smog and global warming, are a salient concern both locally and worldwide. China alone, with 300 million new consumers and possibly twice as many within ten years, could soon exert an environmental impact to rival that of the United States.

Confronting Consumerism begins by conceptualizing the issue as “a problem of political and ecological economy”. It looks at some of its main features and how they interplay with other aspects of society. The book goes on to assess “transnational chains of consumption in the context of economic



Our love of cars is already a pollution issue, and the growth of consumerism in China (right) is also likely to have a huge environmental impact.

globalization”, with emphasis on the export of Western lifestyles and values. A final part describes citizen action and activism that goes far beyond marginal tinkering such as green purchasing and changes in production processes. Although these activities help, they imply that we can continue indefinitely with our runaway consumption practices. We might soon enjoy cars that can achieve 100 miles per gallon but, as previous increases in fuel efficiency have shown, when travel costs are reduced, individuals respond by driving further.

In this final section, the authors advocate more perceptive and discriminating consumption, including “voluntary simplicity”, or the voluntary adoption of less materialist lifestyles. But they refrain for the most part from urging a reining in of consumption. Still less do they propose a no-growth economy. Citizens can expect to keep on keeping



on with their consumption, provided that they undertake fundamental shifts in their consumption patterns. The icon of the industrial age has been the car — two tons of material to get us from here to there, causing a lot of pollution along the way. The icon of the future could prove to be the microchip, which uses so few raw materials that all the world's chips would fit inside a single jumbo jet — and it can obviate the need for us to go from here to there in the first place, by working at home, for instance.

Through 14 chapters by the three editors and seven other authors, the book covers such diverse topics as spending efficiency, responsible shopping, consumer sovereignty (“the consumer knows best”), consumption

AP

NEWSMAKERS

externalities both environmental and social, eco-technologies, quality of life/quantity of livelihood, and the economics of happiness (yes, indeed). All are addressed within such overarching concepts as political economy, social equity, environmental imperatives, and the global north–south divide. Many authors are deliberately provocative, as is warranted by the book's thesis, and most are thoroughly illuminating.

The book's bottom line is to advocate new lifestyles, underpinned by modified value systems. The approach can be summed up by one-liners such as "Goods are good, so more goods must be better", "The best things in life are not things", and "How to advance from 'more is better' to 'enough is best'?"

The book makes an excellent exploration of what could turn out to be one of the front-rank issues of our time. Tackling consumption will not be simple. We have had 10,000 years of believing that more of anything must be better, and have had solid reasons to believe this. Switching to lifestyles that are not

better off but simply better will be a formidable challenge. It raises all manner of questions about the fundamentals of our societies, as this book admirably proclaims. ■

Norman Myers is Honorary Visiting Fellow, Green College, Oxford University, Upper Meadow, Old Road, Oxford OX3 8SZ, UK.

The quest for the Jesuit's bark

The Fever Trail: The Hunt for the Cure for Malaria/The Fever Trail: In Search of the Cure for Malaria by Mark Honigsbaum

Macmillan: 2001. 352 pp. £18.99/

Farrar Straus Giroux: 2002. \$25

Sandra Knapp

If ever there were a story that illustrates the need for high-quality

taxonomy, this is it. The quest for a cure for malaria has occupied many people for many years — and it is still going on, both in the field and in the laboratory. Mark Honigsbaum has brought the story of the search for the first remedy, 'Jesuit's bark' — now known as quinine — to life in an entertaining and readable way, all the while reminding us that malaria is the world's third-biggest killer behind dysentery and tuberculosis.

Some parasitologists estimate that half of the people ever to have lived on Earth have succumbed to this deadly disease. Malaria decimated military might in battle after battle during the eighteenth and early nineteenth centuries, from Britain's attempts to take Spain's New World colonies to critical confrontations with Napoleon in Belgium, where forces were halved through mortality and morbidity. The numbers are hardly less frightening today: there are between

300 million and 500 million cases, and 1.5 million to 2.7 million deaths, each year from malaria, most of them children.

Most of the book is devoted to the travels and travails of the European explorers Richard Spruce, Charles Ledger and Clements Markham. Their stories are woven in and out of the social issues and political climate of Europe in the nineteenth century. Spruce and Ledger in particular were intrepid men, braving much hardship in the search for botanical novelty.

They both became involved in the search for the miracle malaria cure Jesuit's bark, the bark of the *Cinchona* tree, when they were already in South America. Spruce was collecting plants and Ledger was setting up businesses. Botanists in the Andes observed that the forests were being stripped of trees, this potentially renewable resource being extracted rather than used sustainably. Spanish control over the *Cinchona* regions, on the eastern slopes of the Andes, made supplies of the tree limited and expensive. Coupled with the British and Dutch colonization of the tropical regions of Asia, the climate was ripe for some *ex situ* conservation.

Markham and others reasoned that it was better to save this miracle cure by transplanting it into regions where it could be cultivated for the use of all — an argument that surfaces again and again in relation to the destruction of natural resources, even today. The quest was bedevilled by questions of how to tell the species of *Cinchona* apart and which one was the most effective cure — indeed, the complex taxonomy of *Cinchona*

Miracle cure: the *Cinchona* tree was found to yield the antimalarial drug quinine — but it was better to keep mosquitoes at bay in the first place (inset).



was not stabilized until the 1990s. There was also a need for secrecy, because the export of *Cinchona* seeds or bark was illegal, and the Spanish colonies (and the republics that followed) had a monopoly on the raw material for the drug. The story of the theft of *Cinchona* has many parallels to the situation today. To whom do cures 'belong' — the people in the country where the plant grows, or those who discovered the use of the plant? And should we develop the drugs for profit or for philanthropy?

The story that Honigsbaum does not tell is that of the mosquito itself. Only some species of *Anopheles* transmit malaria, and their taxonomy is as complicated as that of *Cinchona*. Controlling the disease by destroying the mosquitoes that transmit it depends on first identifying them, a taxonomic tale not told in this book but perhaps the subject for another. Another problem was that distinguished medical men dismissed as fantastic the idea that mosquitoes transmit a filarial parasite that causes malaria — only dogged perseverance showed how wrong they were.

The development of synthetic drugs such as chloroquine and quinacrine made dependence on *Cinchona* bark a thing of the past, and an understanding of how the disease worked made control seem possible. But here this meant confronting the reality of nature. Malaria parasites evolve drug resistance, mosquitoes evolve pesticide resistance, and a vaccine seems as far away as ever, despite decades of research. The depressing observation by one of Honigsbaum's sources that "pharmaceutical companies are simply not interested in developing drugs for people who cannot afford shoes" shows that the tension between profit and philanthropy is still with us, just as it was in the hunt for *Cinchona* in the nineteenth century. But global warming and the concomitant spread of malaria to Europe and the United States may change all that. Malaria may become a disease of people who have shoes, and the impetus for drug and vaccine development may acquire new urgency.

One of the great riddles of malaria is how *Cinchona*'s properties came to be discovered at all. Malaria did not exist in the New World until the Spanish and Portuguese arrived in the sixteenth century. Malaria is an Old World disease that was initially cured using a New World plant. Honigsbaum doesn't answer this conundrum, but he has written a fascinating book full of thought-provoking ideas. The seemingly arcane and out-of-date story of the hunt for a malaria cure brings home the timely message that knowledge of the world around us, from taxonomy, ecology and medicine, remains critical to our capacity for survival. ■

Sandra Knapp is in the Department of Botany, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

A modern kind of magic

Knowledge is Power: How Magic, the Government and an Apocalyptic Vision Inspired Francis Bacon to Create Modern Science

by John Henry

Icon Books/Totem Books: 2002. 176 pp.

£9.99/\$15

Alan Stewart

The sensational claims of the book's subtitle are highlighted by a cover illustration of Gerrit Dou's mid-seventeenth-century depiction of an astronomer working by candlelight, hinting at some sort of supernatural skulduggery. It must be said that Dou's astronomer doesn't look anything like Francis Bacon, Baron Verulam, and there are bound to be Bacon scholars who would baulk at the idea of magic and apocalyptic visions having much to do with baconian natural philosophy. But John Henry's short study makes an interesting case for looking at the supposed father of modern science in a slightly different light.

Some recent criticism of Bacon has taken issue with his continued pre-eminence in so many fields, and Henry readily admits that the man "made no new discoveries, developed no technical innovations, uncovered no previously hidden laws of nature". But he is convinced of Bacon's importance as "a philosopher of science — perhaps the first one who really mattered". Before Bacon, he argues, "there was no such thing as science in our modern sense of the word". He points to three key factors comprising Bacon's importance: an insistence on experimental method rather than armchair speculation; the notion that a new knowledge of nature should be turned to the practical benefit of mankind; and the championing of inductive over deductive logic. "In a very real sense," he concludes, "Bacon invented modern science."

This is all quite traditional. Where Henry makes the case more pertinently is in his reading of Bacon's motivations: one practical and one spiritual. Following recent biographical emphases, Henry stresses Bacon's status as a career civil servant, saying he "always believed that the fully comprehensive and practically useful science he envisaged could only properly be pursued under the aegis of the state". This vision, most fully explained in the description of Solomon's House in Bacon's fable *New Atlantis*, remained unrealized in his own lifetime. But Bacon was prophetic, Henry argues, in understanding the need for state subsidy of scientific experiment.

Perhaps the most compelling section of the book deals with Bacon's "magic", by which



Dark forces? Gerrit Dou's painting *Astronomer* by Candlelight evokes a world of mystery.

Henry means religion. Here he makes a more convincing case than many for the profoundly religious underpinning of Bacon's philosophical project. The title page of the 1620 *Novum Organum* famously pictures the Straits of Gibraltar, the very edge of the Old World, and the motto "*Multi pertransibunt et augebitur scientia*" ("Many will go to and fro, and knowledge will be increased"). As Henry notes, the line is taken from the Book of Daniel (chapter 12, verse 4), and Daniel is to the Old Testament what Revelations is to the New — here, then, is the apocalyptic vision.

Bacon firmly believed that he was living in the era in which the scriptures predicted that knowledge would increase beyond all recognition. Had not the past decades seen crucial advances in learning, warfare and navigation, in the form (respectively) of the printing press, gunpowder and the magnetic compass, he asked? Part of his *Instauratio Magna* was entitled *Parasceve*, the Greek word for 'preparation', but particularly the day of preparation for the Sabbath, the ultimate Sabbath of the Day of Judgement. "What else can the prophet mean... in speaking about the last times?" Bacon asked rhetorically in his *Refutation of Philosophies* in 1608. "Does he not imply that the passing to and fro or perambulation of the round earth and the increase or multiplication of science were destined to the same age and century?"

Henry expounds his case with the contagious, even bullish, enthusiasm of the committed teacher. The book is carefully pitched at an interested but not necessarily informed readership, and has a useful glossary of the more opaque technical terms and a brief practical guide for further reading. But although *Knowledge is Power* will find its primary audience as an introduction to Bacon, it still has something to offer the seasoned Bacon scholar. ■

Alan Stewart is in the School of English and Humanities, Birkbeck, University of London, London WC1E 7HX, UK.

A progression puzzle

René Bernards and
Robert A. Weinberg

Most, if not all, human tumours develop through a succession of genetic and epigenetic changes that confer increasingly neoplastic (cancer-like) characteristics on cells. Indeed, this multi-step process has been likened to darwinian evolution within the microcosm of living tissues, in which the units of selection are individual cells. A cell that possesses advantageous characteristics (ones that favour survival and proliferation) is selected to become the progenitor of a successor cell population that eventually dominates the tumour mass. A rare variant that arises among the many successor cells will, in turn, initiate the next round of clonal succession. Between six and ten such clonal successions may be required to generate highly malignant human cancer cells.

According to the prevailing reasoning, this process of tumour progression gives rise initially to cells that have acquired the ability to form primary tumour masses of substantial size. Genetic changes, acquired by cells during the initial phases of tumour progression, that provide some type of proliferative advantage enable the cells carrying these mutant alleles to spawn large descendant populations within the primary tumour mass. Among these advantageous phenotypes are the acquisition of constitutive mitogenic signals, the ability to resist growth-inhibiting signals, to avoid programmed cell death (apoptosis) and to induce blood-vessel growth (angiogenesis). Subsequently, individual cells in these large cell populations acquire yet more mutant alleles that enable them to metastasize to seed new colonies at anatomical sites that may be far removed from the primary tumour mass.

This model of tumour progression carries with it a striking

Spreading pattern: the tendency of a breast cancer (yellow) to give rise to secondary tumours may be determined early in its life.

conceptual inconsistency: the genes that specify the final step in tumour progression — metastasis — would not seem to confer increased proliferative benefit at the primary site. That is, there is no reason to think that a metastatic phenotype enables cells to proliferate more effectively within the primary tumour mass, thereby increasing their representation in the overall tumour-cell population. Hence, rare cells in the primary tumour mass that happen to acquire metastatic capability will remain rare. As the success rate of individual cells undertaking metastasis is extraordinarily low, this makes it difficult to imagine how metastasis can ever proceed.

Reasoning like this drives us to consider a quite different mechanistic model: namely, that the tendency to metastasize is largely determined by the identities of mutant alleles that are acquired relatively early during multistep tumorigenesis. It is already apparent that there are several alternative genetic paths that cells can take en route to forming a primary tumour. Thus, a particular phenotype required early in tumorigenesis by an evolving tumour cell can be acquired through the mutation of any one of several alternative growth-controlling genes. We suggest that a subset of the mutant alleles acquired by incipient tumour cells early in tumorigenesis confer not only the selected replicative advantage, but also, later in tumorigenesis, the proclivity to metastasize. This proclivity will become manifest only much later in tumour progression, in the context of yet other mutations that have struck the genomes of descendant cells.

This type of thinking has three implications. First, the tendency of a tumour eventually to metastasize is already preordained by the spectrum of mutations that progenitor cells acquire relatively early in tumorigenesis; that is, some cancers start out 'on the wrong foot'. Second, genes and genetic changes specifically and exclusively involved in orchestrating the process of metastasis do not exist. Instead, the genes for metastasis are largely those that cancer biologists have been studying intensively for a generation: the oncogenes and tumour-suppressor genes. Third, because important components of the genotype of metastasis are already implanted in cells relatively early in tumorigenesis, even relatively small primary tumour cell populations may already have the ability to dispatch metastatic pioneers to distant sites in the body.

Several independent lines of evidence seem to support these ideas. In some small, well-localized primary human breast

Metastasis genes

The prevailing model of tumour progression carries with it a striking conceptual inconsistency.

cancers, individual carcinoma cells are clearly detectable in the bone marrow. Furthermore, DNA-microarray analysis reveals that the gene-expression pattern of metastatic tumour cells is often strikingly similar to that of the cells confined to the primary tumour mass from which they were derived, implying that the dominant cell population in the primary tumour mass is phenotypically and possibly genotypically (almost) identical to the cells in the metastases.

Equally relevant are other studies in which the gene-expression profiles of the dominant populations of breast-cancer cells within a primary tumour mass have been used to predict, with 90% accuracy, whether the tumour will remain localized or whether the patient will experience metastases and disease relapse. Here, once again, the metastatic behaviour of these cancer cells seems to be determined relatively early in tumorigenesis. (The implications for the usefulness of early clinical detection of breast cancer are unsettling.) Finally, several well-studied oncogenes, including *ras* and *myc*, the proliferative powers of which are well documented, can function in certain mouse models of tumorigenesis to enable cancer cells to metastasize.

These ideas have implications for our eventual understanding of the genetic and biochemical bases of metastasis. Some researchers have searched far and wide in the genomes of advanced, highly aggressive tumour cells for the genes responsible for inducing metastatic capability. Perhaps the culprits have been staring us in the face for a long time — the mutant genes that are known to confer darwinian selective advantages early on may be the same genes that, further down the line, empower metastasis. ■

René Bernards is in the Division of Molecular Carcinogenesis and Center for Biomedical Genetics, The Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, The Netherlands.

Robert A. Weinberg is at the Whitehead Institute for Biomedical Research and in the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA.

FURTHER READING

Hanahan, D. & Weinberg, R. A. *Cell* **100**, 57–70 (2000).
van 't Veer, L. J. *et al. Nature* **415**, 530–536 (2002).
Yokota, J. *Carcinogenesis* **21**, 497–503 (2000).
Pozzatti, R. *et al. Science* **232**, 223–227 (1986).

Virus deals anthrax a killer blow

M. J. Rosovitz and Stephen H. Leppla

The threat of antibiotic-resistant bacteria is driving researchers to think up ever more clever ways to tackle infections. An enzyme from a bacterium-killing virus may prove effective against anthrax infections.

Following the deliberate dissemination of spores of *Bacillus anthracis*, the bacterium that causes anthrax, through the US mail last autumn, public-health officials were faced with two major problems: detecting spores in buildings and on exposed individuals, and treating those people thought to be exposed and the few who actually became infected. Testing for spores required samples to be sent to specialized labs to be cultured, and took several days — a painfully long time when the public was insisting on immediate information and action. And thousands of people who were thought to have been exposed were treated with antibiotics, usually ciprofloxacin. Fortunately, nobody treated preventatively in this way became infected; the intervention was effective because the particular strain used in the attack is wholly susceptible to the usual antibiotics. But the situation could have been much worse if the strain had been resistant to antibiotics.

On page 884 of this issue, Schuch and colleagues¹ describe an innovative tool that might help to solve all these problems by rapidly detecting *B. anthracis* spores and curing an infection — even one resistant to standard antibiotics. The group's antimicrobial strategy originated in earlier work by the senior author, Vincent Fischetti, and defines a new class of antibiotics that is distinct from any previously available.

Recent years have seen a growing awareness of the problem of antibiotic-resistant bacteria. Even defined environments such as hospitals present difficulties in dealing with antibiotic-resistant variants of common bacterial infections, suggesting that deliberate, widespread dissemination of spores of an antibiotic-resistant *B. anthracis* strain could cause a far more serious public-health emergency than that which occurred last autumn. These concerns — along with worries about the current anthrax vaccine that have led some in the US military to refuse immunization — have spurred intense research into alternative forms of preventing and treating *B. anthracis* infections.

Effective and more acceptable vaccines are being developed². However, these, like many other vaccines, will require multiple immunizations, and time for protection to build up. To be effective, a vaccine would need to be administered well in advance of an attack. On another tack, recent breakthroughs in understanding the structure and function of



Figure 1 Investigating anthrax. The intentional spread of anthrax spores through the US mail last year has spurred many attempts to develop quicker ways of detecting the spores, and new ways of treating infections. Schuch *et al.*¹ have come up with a simple technique that may meet both aims.

anthrax toxin — one of the two major components the bacterium needs for its virulence — have led to candidate drugs that directly interfere with toxin assembly and function^{3–6}. Such products may protect against the toxic effects of anthrax. But drugs that destroy the bacteria will also be important.

One technique for tackling bacterial infections is to use bacteriophage — viruses that infect and burst open (lyse) specific bacteria. Bacteriophage therapy has had a chequered history, but the recent successful application of a genetic means of selecting viral variants that resist elimination from the human blood circulation⁷ suggests that modern techniques might resurrect the practice. In an elegant variation on this classical approach, Schuch and colleagues¹ have exploited a single key enzyme of bacteriophage γ , a virus often used in clinical labs to identify *B. anthracis*^{8,9}. The authors first purified this enzyme, PlyG, which is a type of lysin; in the normal course of the bacteriophage life cycle, it destroys the bacterial cell wall so as to allow the release of progeny phage. Lysins are typically highly specific for

particular species or strains of bacteria, because their binding regions are directed towards cell-wall carbohydrate structures that vary greatly among species and strains^{10,11}. Accordingly, the PlyG lysin described by Schuch *et al.* was highly effective in killing *B. anthracis*, but did not affect most strains of the closely related species *Bacillus cereus* and *Bacillus thuringiensis*.

Although bacteriophage lysins are made inside bacteria and normally cause lysis from within, the authors showed that PlyG added externally had strong lytic activity towards all the *B. anthracis* strains tested when they were grown on solid or liquid media. Most importantly, Schuch *et al.* also found that when mice were injected with PlyG 15 minutes after being infected with a close relative of *B. anthracis*, about 80% of the animals were rescued from otherwise certain death. An obvious concern is that some of the mice might have died because PlyG-resistant mutants of *B. anthracis* had arisen. However, a controlled test showed that PlyG-resistant mutants were not generated in *B. anthracis* cultures, even when the cultures

ROBERTO BOREA/AP

were treated chemically in a way that increased the percentage of mutants resistant to standard antibiotics.

Schuch *et al.* suggest that the absence of the development of resistance to PlyG is due to the fact that any mutational change to the cell-wall structure that prevents binding to PlyG would kill the bacterium. This targeting of phage lysins to an essential bacterial structure gives them an advantage over the small-molecule antibiotics to which bacteria can become resistant rather easily. One drawback, however, shared by many other candidate antibacterial tools, is that PlyG would have to be administered soon after a person became infected, before lethal levels of anthrax toxin could develop. As illustrated by recent human cases, it is often difficult to diagnose anthrax infections early enough for treatment to be effective.

It is here that PlyG has another possible application. Currently, environmental and clinical samples suspected of being contaminated by anthrax are usually sent to specialized labs for culture. One of the tests used at the Centers for Disease Control and Prevention in the United States involves checking cultured bacteria for sensitivity to bacteriophage γ (refs 1, 8, 9). But all culture methods are slow. Efforts are under way to develop quicker tests using, for example, molecular approaches that require DNA extraction and special instrumentation. However, such sophisticated methods are difficult to deploy outside the lab.

Schuch *et al.* have now shown that PlyG can be used in a simple system to rapidly detect *B. anthracis* spores. Spores are resistant to PlyG-induced lysis, so the authors added the amino acid L-alanine to trigger germination, and then treated the emerging bacteria with PlyG, causing lysis and the release of bacterial components. One of these components is ATP, the main cellular energy store, which the authors detected using a luciferase-luciferin system: the luciferase enzyme degrades luciferin in the presence of ATP, producing light, which can be detected with a hand-held luminometer. This type of system would readily lend itself to field detection of spore contamination, because the equipment could be easily incorporated into a relatively small and simple device. One limitation is that it would not distinguish between virulent and non-virulent strains of *B. anthracis*, because PlyG would lyse the bacteria regardless of whether or not they produced the two factors needed for full virulence — anthrax toxin and a polyglutamate capsule. Presumably, however, the test could be used outside the lab as a rapid first indicator of contamination by *B. anthracis*, and further characterization would follow.

There is still much to be done to develop PlyG into an effective drug. For example, it would probably need to be administered intravenously in a formulation that would

give adequate concentrations in the blood, because that is where the bacteria grow rapidly during the dangerous final stage of infection. Nonetheless, Schuch *et al.*¹ have introduced a potential treatment for anthrax that might be useful either alone or in combination with other therapies. Similarly, their means of detecting *B. anthracis* spores may prove valuable as the basis for an easily deployable test. Given these promising results, we expect to see rapid moves to test the use of bacteriophage lysins in detecting and treating other bacterial infections. ■

M. J. Rosovitz and Stephen H. Leppla are at the National Institute of Dental and Craniofacial Research, and will shortly be moving to the

National Institute of Allergy and Infectious Diseases, Convent Drive MSC 4350, Bethesda, Maryland 20892-4350, USA.
e-mail: sleppla@mail.nih.gov

1. Schuch, R., Nelson, D. & Fischetti, V. A. *Nature* **418**, 884–889 (2002).
2. Leppla, S. H., Robbins, J. B., Schneerson, R. & Shiloach, J. *J. Clin. Invest.* **110**, 141–144 (2002).
3. Mourez, M. *et al.* *Trends Microbiol.* **10**, 287–293 (2002).
4. Chaudry, G. J. *et al.* *Trends Microbiol.* **10**, 58–62 (2002).
5. Pannifer, A. D. *et al.* *Nature* **414**, 229–233 (2001).
6. Maynard, J. A. *et al.* *Nature Biotechnol.* **20**, 597–601 (2002).
7. Biswas, B. *et al.* *Infect. Immun.* **70**, 204–210 (2002).
8. Inglesby, T. V. *et al.* *J. Am. Med. Assoc.* **287**, 2236–2252 (2002).
9. Brown, E. R. & Cherry, W. B. *J. Infect. Dis.* **96**, 34–39 (1955).
10. Lopez, R., Garcia, E., Garcia, P. & Garcia, J. L. *Microb. Drug Resist.* **3**, 199–211 (1997).
11. Loessner, M. J., Kramer, K., Ebel, F. & Scherer, S. *Mol. Microbiol.* **44**, 335–349 (2002).

Atmospheric science

Lasing on a cloudy afternoon

Stephan Borrmann and Joachim Curtius

Brief, high-intensity laser pulses can cause water droplets to emit white light. The technique can potentially be used to analyse the composition of clouds and shed light on how clouds may be affecting climate.

Understanding the role of clouds is one of the major obstacles to successfully modelling Earth's climate. Clouds in the troposphere, the lowest region of Earth's atmosphere, contain water droplets and ice particles ranging in diameter from a few micrometres to tens of millimetres. In *Physical Review Letters*, Catherine Favre and colleagues¹ present a novel way to study small water droplets using powerful lasers. Their technique could be developed into a versatile tool for studying clouds and obtaining information about the size and chemical composition of water droplets.

Droplets in clouds both absorb and scatter radiation from the Sun, and interact with infrared radiation from the Earth. But human activity induces changes in the optical properties of clouds, primarily through the indirect effects of aerosol particles. An increase in anthropogenic aerosol particles, such as soot and sulphate particles, is believed to lead to the formation of clouds consisting of smaller, more numerous droplets, because the same amount of water condenses on a larger number of seed particles. As a result, the cloud appears brighter and scatters solar radiation more efficiently back into space than does a cloud consisting of larger droplets. Anthropogenic aerosols are also thought to influence the lifetimes of clouds and how precipitation develops within them. But to assess these effects quantitatively requires a knowledge of the size distribution and optical properties of cloud droplets.

According to a recent report² by the Intergovernmental Panel on Climate Change (IPCC), the indirect effects of aerosols on climate could be as large as the greenhouse

effect caused by anthropogenic carbon dioxide emissions, although acting in the opposite direction. But there are large uncertainties in estimating indirect aerosol effects for global climate models, and the IPCC report explicitly states that the level of scientific understanding of the underlying processes is "very low". One reason for this uncertainty is the lack of information on cloud–aerosol interactions, which again depend on the chemical composition of the aerosol particles and cloud droplets. And it doesn't stop there: to properly describe man's impact on Earth's climate, the feedback effects of aerosols and clouds on atmospheric temperature and chemistry, and on wind fields and relative humidity, must all be considered³.

Favre *et al.*¹ have found that a water droplet 60 μm in diameter can emit white light in response to a laser flash. The droplet acts as a spherical lens, intensifying the incident radiation 100–200-fold until a small region inside the droplet becomes ionized and forms a plasma (Fig. 1). The plasma is so hot that, as electrons and ions recombine, white light is emitted; the spectrum of the emitted light is that expected for a black-body radiator at a temperature of around 7,000 K. This process is not observed in bulk samples of liquid water under comparable experimental conditions.

To ionize water and create a plasma in the droplet, the intensity of the laser pulse — which lasts just 120 femtoseconds (1 fs = 10^{-15} s) — must be high enough to trigger a multi-photon process involving five or more photons⁴. The laser intensity used in the experiments of Favre *et al.* was around 5×10^{12} W cm⁻². A three-photon process

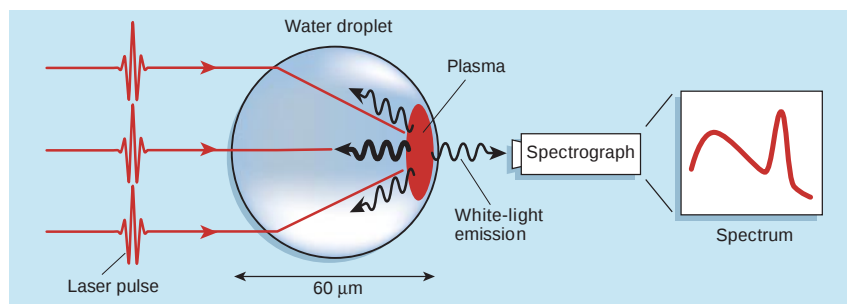


Figure 1 Reach for the clouds. Favre *et al.*¹ have demonstrated that a water droplet exposed to laser pulses emits white light. The droplet acts as a lens, focusing the laser light into a small region which becomes ionized, creating a plasma. As electrons and ions recombine, white light is emitted from the plasma, its distribution peaking in the backward direction. If the emitted light is subjected to spectral analysis, peaks superimposed on its black-body spectrum provide information about the chemicals present inside the droplet.

that was also observed in the droplets was insufficient for plasma generation^{1,5}. The authors also rule out sonoluminescence and another effect called 'self-phase modulation' as causes of the plasma.

The addition of sodium chloride to the water droplets produces a striking feature: superimposed on the white-light spectrum is the distinctive emission line, called D1–D2, caused by excited sodium atoms. This possibility of identifying the 'fingerprints' of chemical components inside an individual water droplet opens up new avenues in experimental cloud physics and chemistry research.

Furthermore, Favre *et al.*¹ found that white light is emitted from a droplet more strongly in some directions than in others. The favoured emission direction is backwards, towards the laser source; emission is next strongest in the forward direction and there are further emission peaks at 150° to this axis. Theoretical calculations support this angular dependence. They also suggest that droplets as small as 22 μm can emit white light if the incident intensity of the laser beam is sufficiently high. This means that a significant fraction of droplets that occur naturally — in clouds, fog, hazes and smog — could be amenable to this laser-induced spectroscopic technique.

A technique currently used to determine the chemical composition of single particles *in situ* is laser ablation mass spectrometry^{6,7}: in this process, particles are transferred into a vacuum and exposed to a powerful, nanosecond-long, ultraviolet laser pulse; this causes the particles to rapidly evaporate, releasing ionized molecular fragments that can be identified by mass spectrometry. However, at present, only particles smaller than around 10 μm can be studied, and although the technique avoids contact with the particles, transfer into the vacuum could change the particle composition. For these reasons, laser ablation mass spectrometry has so far been used mostly to analyse aerosol particles and rarely for larger cloud droplets.

But a hypothetical device designed as a hybrid of a laser ablation mass spectrometer and the femtosecond-laser arrangement used by Favre *et al.*¹ could fill this technological gap and provide information on chemical composition for both large aerosol particles and cloud droplets. Using this device, individual cloud droplets could be sampled and exposed to a femtosecond laser pulse. An appropriate spectrograph could then deliver the chemical-composition information, based on the white-light plasma emission. So far, this has been demonstrated only for sodium dissolved in water droplets, and further studies are needed on different chemical species. However, the limitation of laser ablation mass spectrometry — the fact that only molecular fragments or atoms,

rather than whole molecules, can be detected — would still apply to this instrument.

There is another way to use this kind of technique for studying aspects of the atmosphere^{1,8}. Because the largest fraction of the emitted white light emerges as a nearly collimated beam in the backward direction, it is conceivable that particle size, phase and chemical composition could be determined by remote sensing, by firing laser pulses up into the atmosphere, in a similar way to lidar (light detection and ranging) measurements. Rairoux *et al.*⁸ have already demonstrated the feasibility of plasma analysis for remote gas-phase measurements. The experiments published by Favre *et al.*¹ may well form the cornerstone for the development of innovative instrumentation for cloud and aerosol research.

Stephan Borrmann and Joachim Curtius are at the Institute for Atmospheric Physics, Johannes Gutenberg University of Mainz, Becherweg 21, D-55099 Mainz, and in the Department of Cloud Physics and Chemistry, Max Planck Institute for Chemistry, D-55020 Mainz, Germany. e-mails: borrmann@mail.uni-mainz.de curtius@mail.uni-mainz.de

1. Favre, C. *et al.* *Phys. Rev. Lett.* **89**, 035002 (2002).
2. Houghton, J. T. *et al.* (eds) *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, 2001).
3. Jacobson, M. Z. *J. Geophys. Res.* (in the press).
4. Williams, F., Varma, D. P. & Hillenius, S. *J. Chem. Phys.* **64**, 1549–1554 (1976).
5. Kasparian, J. *et al.* *Phys. Rev. Lett.* **78**, 2952–2955 (1997).
6. Noble, C. A. & Prather, K. A. *Analytical Chemistry of Aerosols* (ed. Spurny, K. R.) 353–376 (Lewis, Boca Raton, Florida, 1999).
7. Murphy, D. M., Thomson, D. S. & Mahoney, M. *J. Science* **282**, 1664–1669 (1998).
8. Rairoux, P. *et al.* *Appl. Phys. B* **71**, 573–580 (2000).

Cell evolution

Mitochondria in hiding

Andrew J. Roger and Jeffrey D. Silberman

The apparent absence of mitochondria in some microbes contributed to the view that they were early offshoots of the eukaryotic line of descent. New evidence tells a different story.

At some point in the history of life, certain bacteria took up residence in other cells, starting a symbiotic relationship that led to their establishment as the mitochondria found within most eukaryotic cells. Eukaryotes are organisms whose cells contain a nucleus, a cytoskeleton, internal membranes and, typically, mitochondria that generate energy by aerobic respiration. For many years, however, it was thought that certain types of protists — unicellular eukaryotes — lack the mitochondrial organelle, supporting the notion that these organisms were primitive offshoots of the eukaryotic lineage that diverged before the mitochondrial symbiosis developed¹.

This view has been challenged by discoveries of the genetic remnants of mitochon-

dria in the nuclei of mitochondrion-lacking protists (reviewed in ref. 2). On page 865 of this issue³, Williams and co-workers go further by showing that the apparent absence of mitochondria in one such group, the Microsporidia, was an illusion. Their cells contain minute mitochondrion-derived organelles containing at least one hallmark mitochondrial protein. These curious structures join a growing list of cryptic, mitochondrion-derived organelles found in eukaryotes that no longer need or are capable of aerobic respiration (for example parasites such as Microsporidia, or protists that live in anoxic habitats).

Microsporidia are intracellular parasites that infect many groups of animals and protists. Early evolutionary trees, based on

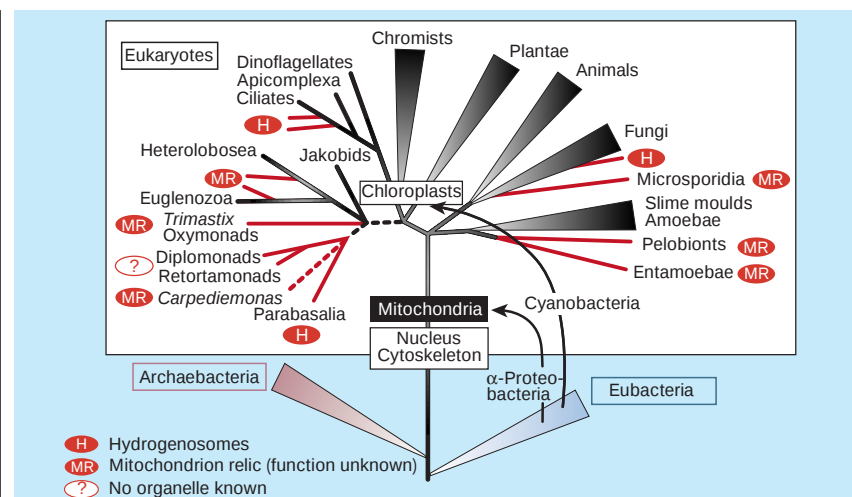


Figure 1 Distribution of mitochondrion-related organelles in eukaryotes. Life is divided into three domains: the eukaryotes, the Archaeobacteria and the Eubacteria. All known eukaryotes probably diverged after the symbiotic development of mitochondria. Eukaryotes with typical aerobic mitochondria are shown in black. Those containing potential mitochondrion-related organelles are shown in red, with an indication of organelle function². Dashed lines indicate uncertainty in branching order. The latest development is the identification³ of mitochondrion relics in Microsporidia, which are parasites related to fungi.

ribosomal RNA sequences, placed them among the most deeply branching eukaryotes⁴, suggesting that they are truly ancient¹. This view persisted until the late 1990s, when sequences of protein-encoding genes began to tell a different story. It now seems that the Microsporidia are not deeply diverging at all, but are specific relatives of the fungi^{5,6} that have adapted to parasitism by economizing in almost every aspect of their biology. They possess the smallest of eukaryotic genomes, and the full genome sequence of one, *Encephalitozoon cuniculi*⁷, indicates that they have dispensed with several metabolic pathways — including some typically found in mitochondria (the Krebs cycle, the mitochondrial electron-transport chain and oxidative phosphorylation). On the other hand, it has also become evident⁷ that Microsporidia have some proteins usually associated with mitochondria, hinting at the existence of an organelle that had escaped detection.

Williams and co-workers³ directly address this issue by examining the expression and localization of the mitochondrial form of a heat-shock protein (mtHsp70) in the human microsporidian parasite *Trachipleistophora hominis*. In other eukaryotes, mtHsp70 is located inside the mitochondrion, where it pulls in and refolds organelle-targeted proteins. Williams *et al.* confirm that *Trachipleistophora* mtHsp70 has a shared ancestry with other microsporidian and fungal forms of the protein; and they show that it possesses functionally important amino-acid residues, and is expressed within the parasite's cells. Most excitingly, they demonstrate that the protein is compartmentalized in numerous 'vesicles' within the microsporidian cell. Immuno-gold

electron microscopy reveals that these structures are tiny, measuring a scant 50 × 90 nm (about one-tenth of the size of most mitochondria), and that, like mitochondria, they are bounded by two membranes. There is little doubt that these 'vesicles' are mitochondrion relics. The organelles are abundant — each cell contains between 7 and 47 of them — hinting at their involvement in an important cellular process.

Clues to what this process may be are provided by the *E. cuniculi* genome sequence⁷. Five proteins of mitochondrial descent that are encoded in this genome, including mtHsp70, are thought to be involved in generating the iron-sulphur (Fe-S) clusters for several proteins that are present in aerobic and anaerobic eukaryotes. If it turns out that these enzymes function within the microsporidian mitochondrion relics, Fe-S cluster biogenesis may be their *raison d'être*. Other puzzling features of these organelles will also require further study, including the apparent absence of a potential (charge differential) across their membranes — a property thought to be required for both mitochondrial protein import⁸ and the synthesis of Fe-S clusters⁹.

Although the cryptic mitochondrion-derived organelles of Microsporidia are surprising, they are not unique. The distinctive hydrogen-producing 'hydrogenosome' organelles of other organisms — parabasalids, certain ciliates and anaerobic fungi — are mitochondria that have been reduced and biochemically refitted with enzymes of anaerobic energy metabolism². More recently, a cryptic organelle (called a mitosome, or crypton), in which mitochondrial chaperonin 60 is found², has been described in the parasitic amoeba *Entamoeba histo-*

lytica. Again, little is known about these organelles except that they are larger and less abundant in cells than are the organelles of *Trachipleistophora*.

Furthermore, putative mitochondrion relics are popping up in all sorts of protists that had been thought to lack mitochondria (Fig. 1). Are these organelles hydrogenosomes, like those of parabasalids, ciliates and fungi, or do they have functions in common with the relic mitochondria of Microsporidia or entamoebae (or both)? Or, as seems most likely, has each organelle retained or acquired a unique suite of functions specific to its host?

Regardless of the answers, it seems clear that mitochondria are rarely, if ever, lost outright. Instead, when their aerobic respiratory functions are no longer necessary, they are reduced and/or remodelled to fulfil some other biochemical function. To date, the only eukaryotes where potential relic mitochondrial organelles have not been identified are two groups of anaerobic protists — the diplomonads, such as *Giardia*², and their retortamonad sisters¹⁰ (Fig. 1). Even so, diplomonads contain at least three proteins of mitochondrial origin that are associated with mitochondria in other eukaryotes¹⁰. Identifying where these proteins reside within diplomonad cells should help to show whether a mitochondrion-derived compartment persists in these organisms.

As we probe further into the nature of the mysterious double-membrane-bounded organelles in anaerobic or parasitic protists, we stand to learn a lot about the functions that mitochondria perform in addition to aerobic respiration. Furthermore, we may have seen only the tip of the iceberg in terms of the range of metabolic capacities that the mitochondrion can acquire when its aerobic functions are no longer needed.

Andrew J. Roger is in the CIAR Program in Evolutionary Biology and Genome Atlantic, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada.
e-mail: aroger@is.dal.ca

Jeffrey D. Silberman is in the Astrobiology Program, Department of Microbiology, Immunology, Molecular Genetics, and the Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California 90095, USA.
e-mail: jsilber@ucla.edu

1. Cavalier-Smith, T. *Nature* **326**, 332–333 (1987).
2. Roger, A. J. *Am. Nat.* **154**, S146–S163 (1999).
3. Williams, B. A. P., Hirt, R. P., Lucocq, J. M. & Embley, T. M. *Nature* **418**, 865–869 (2002).
4. Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. & Woese, C. R. *Nature* **326**, 411–414 (1987).
5. Hirt, R. P. *et al.* *Proc. Natl Acad. Sci. USA* **96**, 580–585 (1999).
6. Keeling, P. J., Luker, M. A. & Palmer, J. D. *Mol. Biol. Evol.* **17**, 23–31 (2000).
7. Katinka, M. D. *et al.* *Nature* **414**, 450–453 (2001).
8. Kispal, G., Csere, P., Prohl, C. & Lill, R. *EMBO J.* **18**, 3981–3989 (1999).
9. Pfanner, N. & Geissler, A. *Nature Rev. Mol. Cell Biol.* **2**, 339–349 (2001).
10. Silberman, J. D. *et al.* *Mol. Biol. Evol.* **19**, 777–786 (2002).

Inside whitecaps

Mark Loewen

Innovative experiments have provided new insights into how bubbles are created by breaking waves. These findings might ultimately lead to more accurate models of global climate.

When ocean waves break, air and sea water mix to form whitecaps. Beneath the surface of the whitecap, a mixture of air and sea water form a violent turbulent flow known as a bubble plume. The plumes generated by typical breaking waves evolve rapidly for approximately 10 seconds as large bubbles rise quickly back to the surface¹. After this time, a relatively diffuse plume of small bubbles persists in the ocean for up to several minutes. Study of the rapidly evolving flows inside these plumes is a challenging task: until now, such studies have not shed any light on the physical mechanisms responsible for bubble creation, nor have they identified the factors that govern bubble size². On page 839 of this issue³, Deane and Stokes report the results of a pioneering study of bubble-creation mechanisms. Their analysis and interpretation of laboratory and field measurements represent a significant step towards unravelling the mystery of bubble creation by breaking waves.

Why should we care about the bubbles generated by breaking waves? Bubbles have a surprisingly important role in many physical, chemical and biological processes occurring at the air–sea interface. Bubble formation increases gas transfer between the air and sea⁴, and rising bubbles scavenge organic material and bacteria from the water column and transport them to the ocean surface⁵. Bubbles are both sources and scatterers of underwater sound⁶, and when they rise back to the surface they burst and eject tiny droplets into the atmosphere. The resulting marine aerosols influence cloud and hurricane dynamics, as well as Earth's radiative balance and biogeochemical cycles⁷.

The foundation for Deane and Stokes's work is the idea that bubbles break up or fragment when the differential pressure forces associated with turbulence exceed the restoring force of surface tension⁸. Thus, in a given turbulent flow, bubbles smaller than a certain size — defined as the Hinze scale — are stable, whereas larger bubbles tend to break up. Deane and Stokes discovered that two primary mechanisms are involved. Larger bubbles are formed by turbulent fragmentation when the air cavity trapped by the overturning wave crest collapses. Smaller bubbles are created by the impact and subsequent splashing that occurs when the toe of the wave crest plunges onto the ocean surface. This mechanism is referred to as 'jet and drop impact'.

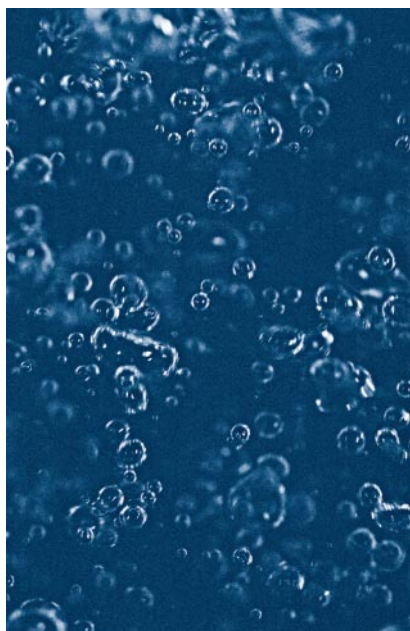


Figure 1 **Bubbling under** — an image from the video recordings made by Deane and Stokes³.

Deane and Stokes used video recordings to measure the size and number of bubbles generated by breaking waves. From these data the authors calculated size distributions showing the number of bubbles per unit volume plotted against bubble radius, and observed a distinct change in the slope of the size distributions at a radius of approximately 1 mm in both laboratory and field data. The video recordings (Fig. 1) revealed that the smallest bubbles that fragmented were about 1 mm in radius. On the basis of these two results and additional analysis, Deane and Stokes conclude that the Hinze scale is equal to approximately 1 mm.

Deane and Stokes's argument that it is turbulent fragmentation that determines the size and number of large bubbles (radius 1 mm and over) is particularly convincing: the observed slope of the size distribution of large bubbles is consistent with an earlier analysis⁹ and with previous measurements of bubble size made beneath the crests of more gently breaking waves¹⁰. The authors acknowledge that their arguments regarding the creation of smaller bubbles by jet and drop impact are more speculative. However, the dimensional analysis and underwater sound data that they present

to support their arguments are consistent and persuasive.

One of the most encouraging aspects of the new work is that it reconciles the authors' laboratory and field measurements of bubble size distributions with their mechanistic model of bubble creation. This not only gives us confidence in the findings but also provides hope that an accurate model of the size and number of bubbles entrained by breaking waves is within reach. Development of such a model is an essential step in accurately assessing the significance of gas transfer by bubbles in the global cycling of gases. As well as using Deane and Stokes's results to provide the shape of the bubble size distribution, a future model will require a reliable method for predicting the occurrence of breaking waves and the volume of air entrained.

The randomness of wave breaking makes it unlikely that theoretical approaches will solve this problem in the near future. But developments in visible⁷ and infrared^{11,12} remote sensing of breaking ocean waves might lead to accurate statistical descriptions of their properties that can be used in lieu of a theoretical model.

Global climate models currently used to predict climate change do not have sufficient resolution to allow direct modelling of small-scale processes such as bubble production by breaking waves. So the effect of these small-scale physical processes on air–sea gas transfer must be incorporated into the models with relatively simple equations. If forecasts of global warming are to be accurate, these equations will have to be based on sound physical principles. The best approach will be to develop equations that combine the statistical information obtained from remote sensing of breaking waves with the fundamental knowledge that stems from experimental studies such as those of Deane and Stokes.

Mark Loewen is in the Department of Civil and Environmental Engineering, Room 220 Civil/Electrical Engineering Building, University of Alberta, Edmonton, Alberta T6G 2G7, Canada. e-mail: mrlloewen@civil.ualberta.ca

1. Lamarre, E. & Melville, W. K. *Nature* **351**, 469–472 (1991).
2. Medwin, H. & Breitz, N. D. *J. Geophys. Res.* **94**, 12751–12759 (1989).
3. Deane, G. B. & Stokes, M. D. *Nature* **418**, 839–844 (2002).
4. Farmer, D. M., McNeil, C. L. & Johnson, B. D. *Nature* **361**, 620–623 (1993).
5. Blanchard, D. C. & Syzdek, L. D. *Science* **170**, 626–628 (1970).
6. Kerman, B. R. *Sea Surface Sound: Natural Mechanisms of Surface-generated Noise in the Ocean* (Kluwer, Norwell, Massachusetts, 1988).
7. Melville, W. K. & Matusov, P. *Nature* **417**, 58–63 (2002).
8. Hinze, J. O. *J. Am. Inst. Chem. Eng.* **1**, 289–295 (1955).
9. Garrett, C., Li, M. & Farmer, D. *J. Phys. Oceanogr.* **30**, 2163–2171 (2000).
10. Loewen, M. R., O'Dor, M. & Skafel, M. *J. Geophys. Res.* **101**, 20759–20769 (1996).
11. Jessup, A. T., Zappa, C. J., Loewen, M. R. & Hesany, V. *Nature* **385**, 52–55 (1997).
12. Zappa, C. J., Asher, W. E. & Jessup, A. T. *J. Geophys. Res.* **106**, 9385–9391 (2001).

Neurobiology

Tuning in by turning off

Ron Hoy

A process that stops crickets from being deafened by their own songs may also explain how they decide whether the songs they hear are their own or another cricket's. It might apply to other senses, and other species, too.

Many animals — among them insects, frogs, birds and mammals such as bats and porpoises — emit ear-splitting calls to attract prospective mates, warn off potential rivals, or echo-locate. Loud is fine for getting a message out. But high power levels could temporarily deafen the songster, and it is just as important to be able to hear responses to the call — whether from rivals with whom the sender might be contesting in an alternating chorus, or from predators attracted to the songster's proclamations. On page 872 of this issue, Poulet and Hedwig¹ reveal a solution to this problem, in the shape of a neurobiological mechanism by which crickets block out the sound of their own songs.

Any animal that produces an intense signal is confronted by two general sensory problems. First, it needs to work out whether its sensory organs are being stimulated by a self-generated or an external signal. Second, if the stimulus is self-generated, the animal's sensory system needs a way of preserving sensitivity to subsequent external signals in the face of potential overstimulation, due to sheer proximity, by its own. Animals have various mechanisms that desensitize their auditory system while they are vocalizing but enable sensitivity to be restored the moment the sound stops. Some animals have sound-sensitive reflexes that temporarily 'load' the eardrums, making them less responsive to sound when stimulated by intense calls². Others deploy neural inhibition to suppress the effect of the overstimulated auditory sensory nerves on the central nervous system (CNS), and so

protect their auditory sensitivity centrally³.

We know rather more about the first type of mechanism than the second, so the findings of Poulet and Hedwig¹ are especially welcome. In a set of technically brilliant experiments, the authors elegantly demonstrate the mechanisms of protective inhibition in the cricket CNS, by recording the activity of a single auditory neuron while experimentally manipulating the insect's sound-producing organs. The inhibition preserves a high level of sensitivity in the auditory system, even while the cricket is producing extremely loud chirps that would otherwise overload its hearing.

Poulet and Hedwig worked with male field crickets, which produce loud mating, courtship and territorial calls. In essence, the authors' experiments required separating the singer from his song. Fortunately, the cricket produces sound pulses by 'stridulation' — rubbing its pair of forewings together. Removing one wing does not impair the insect's ability to make normal stridulatory movements with the other, but the 'sound' of one wing flapping is, of course, silence. Into these one-winged crickets Poulet and Hedwig inserted a microelectrode, in the so-called omega auditory neuron of the CNS; this neuron encodes any cricket songs heard. The authors then monitored responses of this neuron to acoustic signals when the cricket 'sang' its silent song and when it did not.

When the male wasn't (silently) singing, the auditory responsiveness of the omega neuron was perfectly normal, as gauged by the spikes of neuronal activity that were elicited by electronic cricket-like sounds

played through a loudspeaker. But when the cricket was singing, by making stridulating movements with its remaining wing, the omega neuron did not generate spikes; instead, inhibitory synaptic potentials were elicited with every stroke of the wing. There was a one-to-one correlation between stridulation movements and inhibition of the omega neuron.

What is the source of this self-inhibition? It clearly does not come from acoustic stimulation, as a one-winged cricket produces no sound. But it might come from sensory feedback, such as mechanical stretch receptors in the forewing that sense the contraction of stridulatory muscles. To find out, the investigators surgically isolated the wing muscles and associated stretch receptors from the rest of the animal. They then injected the cricket's brain with eserine salicylate, which stimulates the song-pattern-generator network in the CNS ('fictive singing'). From this preparation, it was possible to record action potentials that reflected the generation of the motor pattern of stridulation, even though the motor neurons were disconnected from the stridulatory muscles.

The auditory system and omega neurons in this preparation were acoustically intact, and neuronal spikes could still be produced when sound was emitted from a loudspeaker. But when the preparation was induced to sing fictively, the omega neuron responded to cricket-like acoustic stimuli only with inhibitory synaptic potentials, indicating auditory suppression. Significantly, during the 'silent' gaps between bouts of fictive song, spikes were again elicited from the omega neuron when sounds were played through the loudspeaker. So the inhibition of auditory sensitivity occurs only when the central pattern generator is activated; sensory feedback from moving muscles or joints is unnecessary for auditory suppression.

Poulet and Hedwig interpret this suppression by positing the existence of a 'corollary' discharge from the central pattern generator to auditory interneurons such as

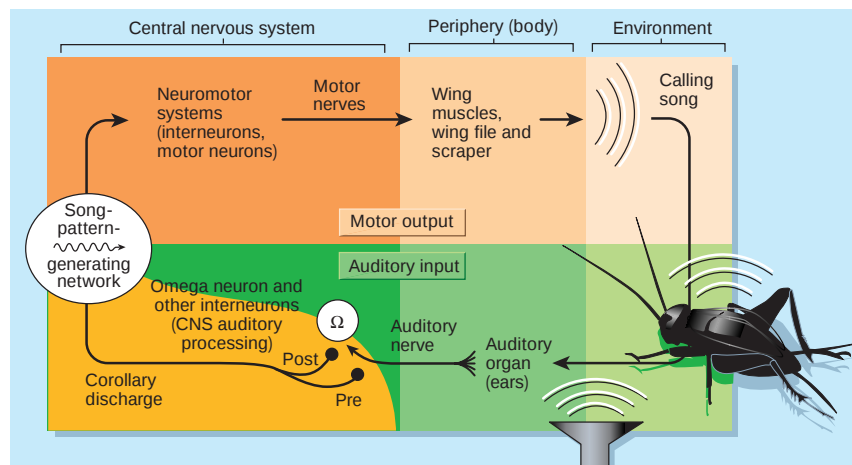


Figure 1 How crickets suppress auditory sensitivity while singing. A cricket sings when it activates the song-pattern-generating network in its central nervous system (left). Once activated, this network has two outputs. One (upper pathway) stimulates the appropriate nerves and muscles to generate the audible calling song. The other output (lower pathway), the 'corollary' discharge, leads to inhibition (by presynaptic and postsynaptic mechanisms) of the omega interneuron, as Poulet and Hedwig show¹. This means that the ability of the omega neuron to be stimulated by auditory input is suppressed as long as the cricket is actively singing. When the insect stops singing, the inhibition of the omega neuron is lifted and its sensitivity to acoustic stimulation (from other crickets or a loudspeaker) returns.

the omega neuron (Fig. 1). The authors even identified two possible sources of this self-inhibition, one arising from the mechanism of presynaptic inhibition, in which the corollary discharge suppresses the excitability of the neurons that stimulate the omega neuron, the other from postsynaptic inhibition (of the omega neuron itself).

This mechanism of dampening auditory sensitivity during singing also deals with the problem of determining whether the auditory stimulation is self-generated or from an outside source: the act of emitting sound simultaneously suppresses the cricket's auditory system as a result of the corollary discharge. This mechanism might also explain how other sensory organs distinguish between self-generated and external stimuli. For example, when we make voluntary eye movements (saccades), the visual scene sweeps across our retinas—but we are usually not aware of this, because our visual system suppresses the self-generated information flow. The underlying mechanism in humans is unclear⁴, but neural inhibition is thought to underlie a similar phenomenon in insects⁵. A rather more familiar example is that of when, if we try to tickle ourselves, we don't collapse

in paroxysms of uncomfortable laughter; the sensation only arises when it's another person doing the tickling. We can't explain this yet in terms of neural processes, but inhibition initiated by corollary discharge is plausible.

Another important function of the auditory system is to identify the location of a singer in its surroundings. The omega neuron has been implicated in how crickets locate and focus their attention on specific sounds⁶, and the newly revealed inhibitory mechanisms, particularly presynaptic inhibition, could also be useful for computing the source of a sound or for focusing attention. So insect systems such as this continue to provide intriguing and salient models for sensory processing, especially in the context of communication. ■

Ron Hoy is in the Department of Neurobiology and Behavior, 215 Mudd Hall, Cornell University, Ithaca, New York 14853, USA.
e-mail: rhh3@cornell.edu

1. Poulet, J. F. A. & Hedwig, B. *Nature* **418**, 872–876 (2002).
2. Narins, P. M. J. *J. Acoust. Soc. Am.* **91**, 3551–3557 (1992).
3. Suga, N. & Schlegel, P. *Science* **177**, 82–84 (1972).
4. Palmer, S. E. *Vision Science* (MIT Press, Cambridge, Massachusetts, 1999).
5. Zaretsky, M. & Rowell, C. H. F. *Nature* **280**, 583–585 (1979).
6. Pollack, G. S. *J. Neurosci.* **8**, 2635–2639 (1988).

Planetary science

Smog report

Robert E. Samuelson

The hazy atmosphere surrounding Titan, Saturn's largest satellite, changes with the seasons. New theoretical work suggests how the motion of smog particles can account for the curious features of the haze.

During the course of a year, planetary atmospheres are, as a rule, heated by the Sun more at the equator than at the poles. This creates a temperature difference between equator and pole, and an atmosphere will attempt to reduce this difference by transporting heat dynamically between them. On Earth this happens mainly through horizontal waves in the atmosphere that produce the cyclone–anticyclone weather patterns familiar to viewers of TV weather reports. The typical diameter for such a cyclone is 1,500 km. But on Titan a typical cyclone would exceed the dimensions of the planet¹, so there must be a different circulation pattern at work.

The preferred alternative explanation for Titan is a slow, meridional circulation in which warm air rises, flows to colder latitudes where it descends, then returns along or near the planet's surface to the original latitude (Fig. 1). On page 853 of this issue, Rannou *et al.*² demonstrate this circulation pattern quantitatively by using Titan's haze, or smog, to trace the motion. An important outcome of their work is an explanation for a curious feature of Titan's haze—a detached,

spherical shell of material separated from the lower main haze by about 100 km (ref. 3).

Although a satellite of Saturn, Titan is a

planet in its own right, with an equatorial radius of 2,575 km. Most of our knowledge of this planet comes from the Voyager spacecraft that flew past Titan in 1980 (Fig. 2). It is one of only two satellites in the Solar System known to have an atmosphere (the other is Neptune's satellite, Triton), and its atmospheric surface pressure is about 1.5 times that of Earth. The atmospheric temperature decreases in the troposphere (the innermost atmospheric layer) from 94 K at the surface to 71 K at an altitude of 42 km (the 'tropopause'), and then increases again to about 175 K at 300 km, the top of the stratosphere.

Titan's atmosphere is mostly molecular nitrogen, with a small percentage of methane. Many hundreds of kilometres above, ultraviolet solar radiation and high-energy electrons (accelerated by Saturn's strong magnetic field) impact on the atmosphere with sufficient energy to fragment both molecular nitrogen and methane. These fragments are highly reactive, and quickly combine into hydrogen cyanide, a nitrile, and the hydrocarbons acetylene, ethylene and methyl acetylene⁴. The organic compounds diffuse and circulate downwards, where further chemical reactions produce more hydrocarbons and nitriles. Almost all organics produced become abundant enough to condense in the cold, lower stratosphere; from there they precipitate into the troposphere, where they become seed nuclei for methane hailstones. These hailstones grow rapidly in Titan's supersaturated troposphere, then rain down on the planet's inhospitable surface.

Alongside the organic volatiles, and central players in Rannou and colleagues' study², are the more complex, non-volatile macromolecules that are formed near the 400-km level⁵. They grow quickly by coagulation, forming a smoggy haze in Titan's

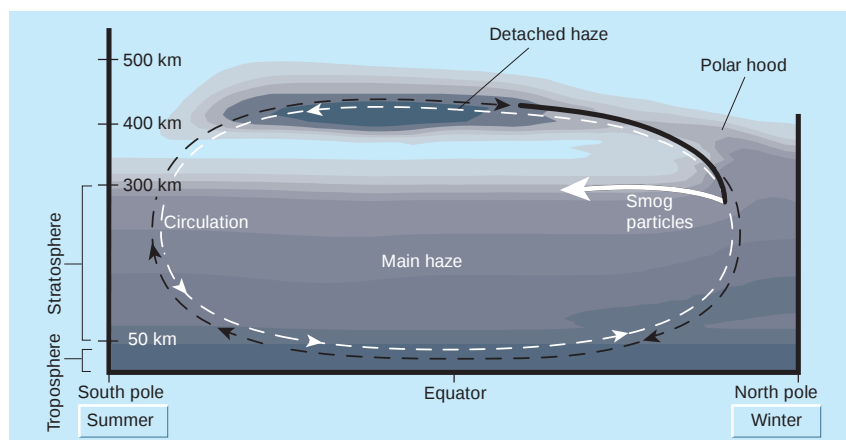


Figure 1 **Going with the flow.** Heat differences drive the circulation of Titan's atmosphere from hot, summer hemisphere to cold, winter hemisphere (dashed black arrows). Smog particles are known to form a 'detached haze' above the planet's atmosphere—a feature that can be explained by a new model from Rannou *et al.*² which shows how, at the change of the seasons and the reversal of circulation (dashed white arrows), the smog particles begin their new traverse at a lower altitude (solid white arrow), leaving a gap below the detached haze.

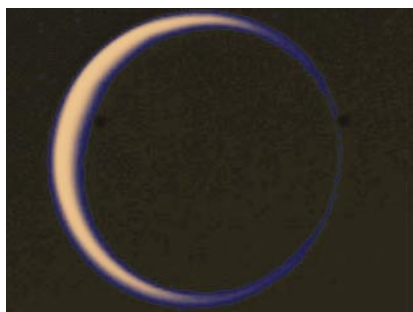


Figure 2 **Ring around Titan.** The planet's thick atmosphere scatters light from the Sun in this image taken by the Voyager 2 spacecraft.

stratosphere. It is this smog that Rannou *et al.* use as a tracer in their model, which consists of a two-dimensional general circulation model, similar to the three-dimensional models used to predict the weather on Earth, combined with a theoretical microphysics model that is used to compute how smog particles grow with time.

The smog and organic gases on Titan share a common characteristic — both are born at altitudes higher than where they die. The volatile organics are produced in the upper stratosphere and above, and are lost through condensation and precipitation from the cold, lower stratosphere below. Smog particles are formed near the 400-km level, diffuse downward and are eventually swept up and removed by droplets of condensing volatiles, such as ethane and acetylene, that are forming near the tropopause. So both are more abundant above the stratosphere than below. It follows that the expected circulation pattern, moving from summer to winter hemisphere at high altitudes and then returning at low altitudes, brings material enriched in both smog and volatiles down into the stratosphere near the winter pole and takes up impoverished material near the summer pole (Fig. 1). This effect has already been demonstrated for the volatile organics⁶, and is now shown quantitatively for the smog by Rannou *et al.*²

Detailed results are shown on page 854 in Fig. 1 of the paper, and are represented in Fig. 1 here. The feature that stands out is the previously mysterious detached haze, for which there is now a simple explanation. Individual smog particles grow into small spheres as they are transported horizontally from the summer to the winter hemisphere. During this period the particles remain small and fall only slowly, tracing out an extended shell. They then circulate downward at high latitudes, building up the 'polar hood' seen in Voyager spacecraft images⁷. When the seasons reverse — Titan's year lasts 30 Earth years — so does the circulation. But by this time the smog particles have fallen to about 100 km below the extended shell, and they make their return trip at a lower altitude, leaving a gap between them and the shell. The individual

particles are now growing rapidly by coagulation, and the speed at which they are falling towards the planetary surface increases such that subsequent changes of season cannot create more, lower, gaps in the haze.

Rannou and colleagues' model² is a first attempt at a quantitative analysis of how dynamics affect smog particles as they grow, but there are several inconsistencies with what is known from observations. For example, the observed winter polar vortex is not taken into account; the predicted accumulation of smog at the summer pole doesn't occur; the extended shell is computed to be too high by 50–100 km, and its height variation with latitude is not quite correct. Finally, no allowance is made for probable cleansing of the troposphere by methane rain.

In spite of these inadequacies, the model

Conservation biology

Openness in management

William J. Sutherland

Should conservation strategy concentrate on intensive management involving practices such as mowing, or should the aim be to protect wilderness? Studies of past ecological conditions can inform that debate.

There is a striking difference between conservation strategies in western Europe and North America, and one that is rarely made explicit. In western Europe, the emphasis is on intensive reserve management, compared with North America and much of the rest of the world where the main aim is to protect wilderness, especially dense woodland. These approaches are based upon an impression of the past — in particular, that closed forest was often the natural state. Writing in *Biological Conservation*, J.-C. Svenning¹ has re-evaluated these impressions in the European context. He has looked at the literature on the pollen record, and records of vertebrate and invertebrate occurrences, to show that the pre-agricultural European landscape consisted of a combination of extensive open areas combined with closed forest. When the past is taken as the template for conservation strategies, research such as this is invaluable.

A guiding principle for conservation in North America is that much of the eastern side of the continent was once forest, but that European settlers cleared large tracts of it. The aim of conservationists, then, is to protect and recreate woodlands to restore the original state. By contrast, in western Europe the main principle stems from the long history of human modification of the landscape, starting with Neolithic farmers who either managed woods intensively by repeated cutting and felling, or cleared them altogether. The species now present in the cleared areas are those that survive or

is a valuable first effort. The Cassini space probe, fresh from its successful fly-past of Jupiter, is now travelling towards Titan. When it approaches that planet in 2004, Cassini will provide exciting new data with which to test and improve the atmospheric models of the future.

Robert E. Samuelson is in the Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA.

e-mail: u3res@lepvax.gsfc.nasa.gov

1. Leovy, C. B. & Pollack, J. B. *Icarus* **19**, 195–201 (1973).
2. Rannou, P., Hourdin, F. & McKay, C. P. *Nature* **418**, 853–856 (2002).
3. Rages, K. A. & Pollack, J. B. *Icarus* **55**, 50–62 (1983).
4. Yung, Y. L., Allen, M. & Pinto, J. P. *Astrophys. J. Suppl.* **55**, 465–506 (1984).
5. Cabane, M., Chassefière, E. & Israel, G. *Icarus* **96**, 176–189 (1992).
6. Lebonnois, S., Toubanc, D., Hourdin, F. & Rannou, P. *Icarus* **152**, 384–406 (2001).
7. Smith, B. *et al. Science* **212**, 163–191 (1981).



Figure 1 **The Oostvaardersplassen — an experiment in maintaining a mosaic of habitats by natural means.**

flourish in open habitats, and their well-being depends on the continuation of farming practices such as mowing, stock grazing or scrub clearance. For example, all of the threatened grassland plant species in the Lorraine region in France require some management to prevent the encroachment of scrub, although the optimal degree of cutting or grazing varies between species².

But how has the landscape of Europe changed? In pre-agricultural times, was it largely dense forest as is often assumed? Vera³ has provided an alternative interpretation, in which there was a shifting mosaic of grassland, scrub and woodland maintained by

herbivores such as deer, moose, wild boar, bison and the wild antecedents of domesticated cattle and horses. His arguments include the following points: hazel is abundant in the pollen record yet does not flower within dense woods; oak rarely regenerates in dense woods but does so within open habitats; and many other plants and animal species that are currently present depend on open habitats.

To examine the generality of Vera's ideas, Svenning¹ reviews the evidence from pollen data to assess vegetation structure in western Europe during previous interglacials and since the last Ice Age but before the advent of agriculture. He concludes that there was a complex pattern and that the extent of open habitat varied with physical conditions. The evidence suggests that there was frequent open habitat as a result of herbivore grazing or fire on flood plains, infertile soils and chalk. Open areas were abundant in dry and warm areas. On fertile uplands there was predominantly closed forest with some, more open, patches maintained by herbivores.

Pollen analysis has a range of associated biases — not least, grazed plants do not produce pollen. So Svenning also looked at the literature on vertebrate remains. Thus, the existence of lynx and woodmouse implies the presence of forest, whereas spotted hyena and field voles are indicative of much more open habitat. Similarly for birds: the presence of capercaillie indicates woodland; red kite, magpie and golden eagle suggest open habitat. Although some of the animal communities characteristic of open habitats were found in coastal regions, flood plains or arid regions, many others were not. So the message to emerge from this analysis is the same as that from the pollen data: during previous interglacials, and in the pre-agricultural present interglacial, there were clearly large areas of open or semi-open habitat as well as more closed forest. This message is also supported by the analysis of invertebrate remains.

For North America, it is commonly thought that before the arrival of European settlers there was almost continuous forest between the Atlantic and the Great Plains, with grassland habitat and grassland species arriving later. However, there are several lines of evidence suggesting that there have always been considerable areas of more open habitat. For instance, three species of bird that favour grassland each have separate eastern subspecies, implying that they have long occurred in isolation, and there are several endemic species of grassland plant⁴. Moreover, the observations of early Europeans in North America, and the pollen, bird and mammal records, indicate that open-habitat communities existed⁴. In both Europe and North America, then, 'wilderness' does not necessarily mean continuous forest.

How might a change in historical perspective within Europe change conservation

practices? Insight comes from the Oostvaardersplassen in the Netherlands (Fig. 1), a 5,600-hectare reserve created in 1974–78 in the lowest section of a recently created polder. The philosophy has been to allow water levels to fluctuate within wide limits, to encourage natural processes and to reinstate the historical herbivore community. As well as the roe deer already present, 150 more red deer and some beavers were introduced. To replace extinct species the introductions included 200 Polish ponies — descendants of wild tarpan, the original European wild horse — and 300 heck, a cow bred to resemble the extinct auroch. Unexpectedly, 60,000 grey-lag geese have appeared in this site each autumn and their grazing has been a major factor in maintaining open habitat over much of the wetland. This experiment is still in its early stages and has its critics, but it has had many ecological successes, including some remarkable populations of scarce birds. The herbivorous mammals and birds have created a mosaic of habitats, including many open areas. Overall, with the lack of obvious human intervention, the impression is one of wilderness.

The main message is that natural habitats included open areas, upon which many species are dependent. The options are to allow such open habitats to disappear (as is often happening in North America), to

restore natural processes (the Oostvaardersplassen model) or to manage with the use of farming techniques (the western European approach). Each option has its own strengths when applied in the appropriate location, but it is likely that everyone could learn from the approaches of others. Conservationists in North America could think more seriously about the importance of biodiversity within open habitats, such as farmland⁵, and ways of managing them⁴. Those in western Europe could look more towards a North American model of nature conservation in which there is a greater distinction between farmed and unfarmed landscapes. The lessons from restoring natural grazing regimes are probably of wide applicability. Detailed accounts of the ecological history, as described by Svenning¹, can greatly inform thinking on these issues. ■

William J. Sutherland is at the Centre for Ecology, Evolution and Conservation, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

e-mail: w.sutherland@uea.ac.uk

1. Svenning, J.-C. *Biol. Conserv.* **104**, 133–148 (2002).

2. Muller, S. *Biodiversity Conserv.* **11**, 1173–1184 (2002).

3. Vera, F. W. M. *Grazing Ecology and Forest History* (CABI, Wallingford, 2000).

4. Askins, R. *Restoring North American Birds: Lessons from Landscape Ecology* (Yale Univ. Press, New Haven, 2002).

5. Freemark, K. E., Boutin, C. & Keddy, C. J. *Conserv. Biol.* **16**, 399–412 (2002).

Immunology

A block at the toll gate

Tak W. Mak and Wen-Chen Yeh

Inflammatory immune responses are crucial to our body's well-being, but too strong a response can be harmful. A protein that keeps inflammation under control has now been discovered.

Inflammation is often thought of as a bad thing, because it is characterized by swelling and pain in affected tissues. Yet we couldn't do without this arm of our immune defences. Inflammation can be triggered when proteins known as Toll-like receptors (TLRs) detect specific patterns of molecules derived from bacteria or viruses¹. The TLRs, which are expressed on the surface of cells of the immune system, are at the front line in the fight against invading microorganisms. In response to TLR engagement, an intracellular signalling cascade is set off that leads to the production of intercellular messenger proteins called cytokines. Some of these, such as interleukin-1, initiate inflammation by increasing blood flow to infected areas and recruiting other immune cells. However, cytokine production must be tightly controlled: excessive production leads to amplified inflammatory responses and devastating illnesses such as acute septic shock and chronic arthritis. The molecules

that signal from TLRs to cytokines have thus been eagerly sought, in the hope of discovering a means of controlling inflammation.

In an important paper that appeared recently in *Cell*, Kobayashi and colleagues² describe one such molecule. The authors generated mice with a defective gene for the IRAK-M protein (which stands for 'interleukin-1-receptor-associated kinase M') — a member of a family of four proteins associated with the intracellular signalling pathway leading to inflammation. Unlike mice lacking other IRAK proteins, mice without IRAK-M show increased inflammatory responses to bacterial infections. This protein is therefore a negative regulator of inflammatory signalling.

On stimulation by specific microbial patterns, different TLRs trigger a common intracellular signalling cascade involving the adapter proteins MyD88 and TRAF6, and the IRAK proteins. This pathway eventually leads to the activation of cytokine-encoding

genes. IRAKs contain regions that allow them to interact with other proteins, as well as domains that are thought to work as kinase enzymes (adding phosphate groups to target proteins); they have generally been regarded as positive signal transducers. Indeed, gene-targeted mice lacking IRAK-1 or IRAK-4 show defective responses to TLR engagement³⁻⁵.

The findings of Kobayashi *et al.*² are surprising because mice without IRAK-M show responses to TLR stimulation that are greater than normal. In addition, cells lacking IRAK-M produce larger amounts of pro-inflammatory cytokines in response to various TLR-activating molecular patterns. Moreover, the activation of the gene-transcription factor NF- κ B (required to switch on the pro-inflammatory cytokine genes) and stress kinases (which contribute to inflammation) is increased or accelerated. Kobayashi *et al.* also show that the lack of IRAK-M enhances signalling through TLR9 (which is triggered by certain patterns in bacterial DNA) more than through TLR4 (which responds to lipopolysaccharide, a major component of the cell walls of many bacteria) and TLR3 (stimulated by viral double-stranded RNAs). This observation is consistent with previous reports that, in addition to the common signalling cascade, individual TLRs might also use unique pathways⁶.

Finally, Kobayashi *et al.* found that cells without IRAK-M respond more strongly than usual to interleukin-1. This cytokine can amplify inflammatory responses *in vivo*, and the interleukin-1 receptor has a similar protein structure to those of the TLRs. These results indicate that IRAK-M negatively regulates signalling both through TLRs and through the interleukin-1 receptor, and imply that modulating IRAK-M activity might affect both the initiation and the amplification of inflammatory signals.

An excessive inflammatory response to lipopolysaccharide can induce fatal septic shock, a reaction that can be prevented by a phenomenon called lipopolysaccharide tolerance. Here, cells that have been previously exposed to lipopolysaccharide (or other TLR stimulants) become less responsive to subsequent challenges with this bacterial cell-wall component. The phenomenon could be due to the exhaustion of positive signalling components^{7,8}. Alternatively, a dominant negative-feedback signal might be involved. Indeed, Kobayashi *et al.* have found that cells lacking IRAK-M show perturbed lipopolysaccharide tolerance, although it remains unknown what happens in the whole animal.

How, then, might IRAK-M work? A current model of IRAK-mediated inflammatory signalling is shown in Fig. 1. After a TLR is stimulated, IRAK-1 is recruited to the receptor's intracellular tail by MyD88 and

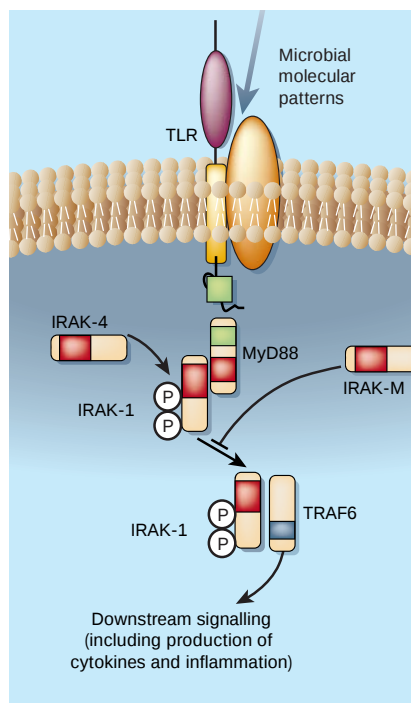


Figure 1 Inflammatory responses. This model of signalling through Toll-like receptors (TLRs) centres on proteins of the IRAK family. After recognizing specific microbial-pathogen-associated molecular patterns, TLRs bind to the intracellular adapter protein MyD88. This protein recruits IRAK-1 to the receptor signalling complex; IRAK-1 is then phosphorylated (circled 'P') and activated by itself, and potentially by IRAK-4 as well. Activated IRAK-1 then leaves the MyD88-TLR complex and associates temporarily with TRAF6, thereby activating downstream signalling. Kobayashi *et al.*² propose that IRAK-M, a family member that does not show phosphorylation activity, negatively regulates the signalling cascade, potentially preventing IRAK-1 from moving to TRAF6.

is activated, presumably by phosphorylation. For the signal to progress beyond this point, IRAK-1 must then leave the MyD88-TLR complex and transiently associate with TRAF6.

IRAK-M can also interact with MyD88 and TRAF6 (ref. 9), but how it inhibits signalling is a matter of speculation. The expression of IRAK-M is very low in resting cells, but high after treatment with lipopolysaccharide. The protein has no kinase activity and might simply compete with the active IRAK-1 for association with MyD88 or TRAF6. However, Kobayashi *et al.* show that IRAK-M can enhance the interaction between overexpressed IRAK-1 and MyD88, and that excess IRAK-M inhibits the interaction between endogenous IRAK-1 and TRAF6. So perhaps, the authors propose, IRAK-M 'traps' IRAK-1 in the MyD88-TLR complex so that it cannot

interact with TRAF6, thereby derailing downstream signalling.

It is also unclear how the phosphorylation, and hence the activation status, of IRAK-1 would be affected by a lack of IRAK-M. The stimulation of TLRs or the interleukin-1 receptor leads to rapid degradation of phosphorylated IRAK-1, depleting the intracellular pool of this protein available for signalling — presumably to prevent too much signalling. Perhaps IRAK-M is involved here; it would be interesting to examine the fate of IRAK-1 in stimulated cells lacking IRAK-M.

Another wild card is IRAK-4, which associates with MyD88 and IRAK-1 and might activate the latter protein^{5,10}. What are the functional relationships between the IRAK proteins? For example, are there any changes in IRAK-4's expression or activity in cells lacking IRAK-M? Another possibility is that IRAK-M — and perhaps IRAK-2, which is induced by signalling through TLRs but has no kinase activity — is a substrate for IRAK-1 or IRAK-4 and is 'activated' by phosphorylation. IRAK-M might then inhibit these kinases in a feedback mechanism that helps to damp down inflammatory signalling.

Whatever the answer, Kobayashi and colleagues' important paper revamps our perception of how IRAK proteins work, in that they can no longer be regarded as functionally redundant. Antagonism of a protein's function by a member of the same family is not unprecedented — witness, for example, the various members of the Bcl-2 family with their different roles in cell death. In immunity, the interactions between IRAK proteins are essential for balancing inflammatory responses. Further studies are required to define these interactions clearly — an important task, given the number of crippling human diseases that are caused by excessive inflammation. Meanwhile, the discovery of a protein that dampens inflammatory responses will no doubt provide food for thought for drug discoverers.

Tak W. Mak and Wen-Chen Yeh are at the Advanced Medical Discovery Institute of the University Health Network, Toronto, and the Department of Medical Biophysics, University of Toronto, 620 University Avenue, Suite 706, Toronto, Ontario, Canada M5G 2C1.

e-mail: tmak@uhnres.utoronto.ca
wyeh@uhnres.utoronto.edu

1. Akira, S., Takeda, K. & Kaisho, T. *Nature Immunol.* **2**, 675–680 (2001).
2. Kobayashi, K. *et al.* *Cell* **110**, 191–202 (2002).
3. Kanakaraj, P. *et al.* *J. Exp. Med.* **187**, 2073–2079 (1998).
4. Thomas, J. A. *et al.* *J. Immunol.* **163**, 978–984 (1999).
5. Suzuki, N. *et al.* *Nature* **416**, 750–756 (2002).
6. O'Neill, L. A. *Trends Immunol.* **23**, 296–300 (2002).
7. Nomura, F. *et al.* *J. Immunol.* **164**, 3476–3479 (2000).
8. Jacinto, R., Hartung, T., McCall, C. & Li, L. *J. Immunol.* **168**, 6136–6141 (2002).
9. Wesche, H. *et al.* *J. Biol. Chem.* **274**, 19403–19410 (1999).
10. Li, S., Strelow, A., Fontana, E. J. & Wesche, H. *Proc. Natl Acad. Sci. USA* **99**, 5567–5572 (2002).

Rapid renewal of auditory hair bundles

The recovery time after noise-induced hearing loss is in step with a molecular treadmill.

Stereocilia, also known as hair bundles, are mechanosensitive organelles of the sensory hair cells of the inner ear that can detect displacements on a nanometre scale and are supported by a rigid, dense core of actin filaments. Here we show that these actin-filament arrays are continuously remodelled by the addition of actin monomers to the stereocilium tips, and that the entire core of the stereocilium is renewed every 48 hours. This unexpected dynamic feature of stereocilia will help our understanding of how auditory sensory function develops and is maintained.

Stereocilia develop before or immediately after birth according to a hair-cell developmental programme that generates a 'staircase' bundle of stereocilia (Fig. 1a) of precise number and height¹, and which grow at about 0.5 μm per day¹. However, the molecular mechanisms by which stereocilia are formed and maintained are poorly defined^{1,2}.

To determine the locus of actin polymerization during stereocilium development and the degree of actin turnover after the stereocilia are fully formed, we exploited the preferential localization of the β -isoform of actin in stereocilia³ (Fig. 1b). We used stereocilia from hair cells of the organ of Corti from newborn rats, which can be maintained in culture for up to 15 days^{4,5}. During this time they develop a complex structural⁴ and functional⁵ phenotype that is similar to that of their *in vivo* counterpart.

We monitored the transient expression of β -actin and its incorporation into actin filaments in stereocilia after transfecting these cultures with a plasmid DNA construct encoding β -actin and enhanced green fluorescent protein (actin-GFP). We used a 'gene-gun' system to transfect segments from the middle turn of the organ of Corti that were dissected at postnatal day 1–3 and maintained in culture. We fixed cultures at various times after transfection and counterstained the total actin with rhodamine-phalloidin before viewing samples through a confocal microscope.

Hair cells rapidly and progressively incorporated actin-GFP, which became visible at the tips of stereocilia within 4 h of transfection and then progressed towards the base. In young (postnatal days 3–5) cultures ($n=27$), actin-GFP was evident along the entire length of the stereocilia as early as 6 h after transfection (Fig. 1c). From this observation, we calculate that actin filaments polymerize at a rate that is about 50 times faster than was previously reported¹.

In our 10–15-day-old cultures, the presumably mature bundles^{6,7} continued to

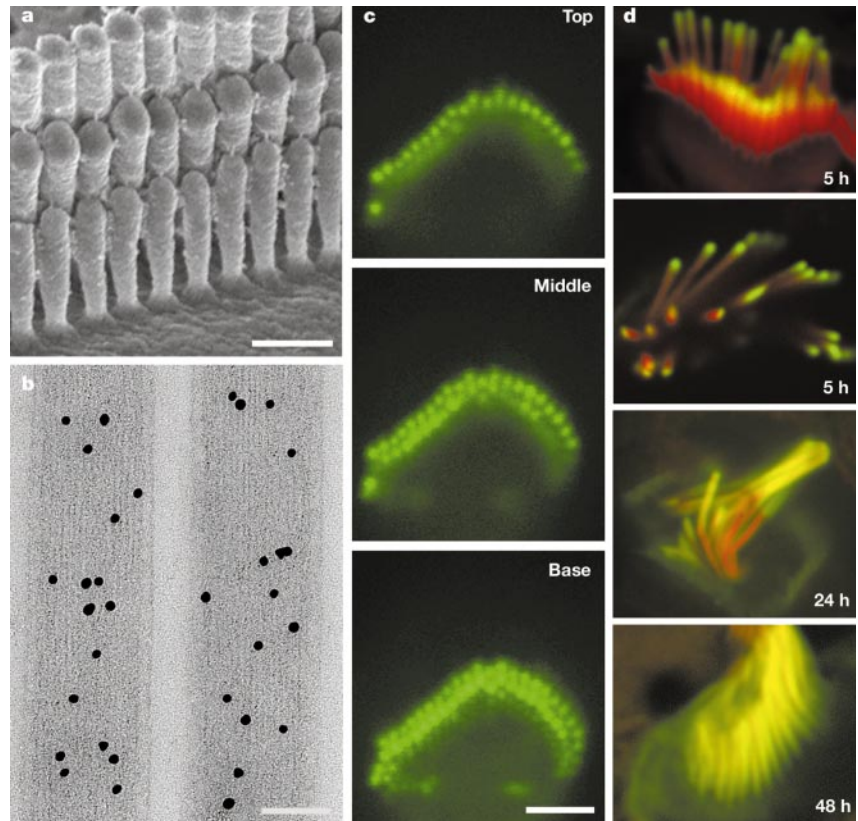


Figure 1 Continuous renewal of actin in auditory stereocilia. **a, b**, Scanning (**a**) and thin-section (**b**) electron micrographs of stereocilia, processed as described¹⁰. Immunogold labelling (**b**) reveals the homogeneous distribution of native β -actin along the semicrystalline array of actin filaments in two stereocilia. **c, d**, Fluorescence confocal images of actin-GFP incorporation (green) in stereocilia of rat hair cells from cultured organ of Corti. **c**, Stereocilia bundle from a postnatal-day-5 culture, 5 h after transfection; cross-sections of the top, middle and base of the bundle show incorporation along the length of the three rows of stereocilia. **d**, Double-labelled images of stereocilia from 10–15-day-old cultures at the indicated times after transfection. Actin filaments are stained red; regions of overlap between red-stained actin and newly incorporated β -actin-GFP appear yellow. Incorporation begins at the tip, has progressed halfway down the stereocilium after 24 h, and has reached the base after 48 h. The progression rate is 2.5 ± 0.18 μm per day ($n=127$; calculated from the length of the GFP-labelled region in the tallest row of stereocilia from hair cells ($n=30$) transfected for 24 or 48 h). Scale bars, 1 μm (**a**), 0.1 μm (**b**) and 2 μm (**c, d**).

incorporate actin-GFP from tip to base until the entire length of the stereocilium was labelled (after 48 h; Fig. 1d). Incorporation was progressive and uniform in these fully developed stereocilia, indicating that each entire actin-filament bundle treadmills towards the base at a rate of 2.5 μm per day (Fig. 1) as new actin monomers are incorporated into the filaments at the tip.

This rapid, systematic incorporation of actin monomers at the tips of stereocilia in mature hair bundles is evidence of a previously unknown, continuous renewal process that is analogous to the migration and renewal of the photoreceptor discs in the retina⁸. The turnover rate of 48 h for this process in mature bundles is of the same order as the recovery time from noise-induced temporary hearing loss⁹, indicating

that the core structure of stereocilia may play an unforeseen part in this recovery. Our discovery that the renewal of stereocilia is dynamic may be of value in the investigation of conditions associated with malformation or disruption of stereocilia.

Mark E. Schneider, Inna A. Belyantseva, Ricardo B. Azevedo, Bechara Kachar

Section on Structural Cell Biology, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892-4027, USA
e-mail: kacharb@nidcd.nih.gov

1. Tilney, L. G., Tilney, M. S. & DeRosier, D. J. *Annu. Rev. Cell Biol.* **8**, 257–274 (1992).
2. Pickles, J. O., Billieux-Hawkins, D. A. & Rouse, G. W. *Hearing Res.* **101**, 45–54 (1996).
3. Hofer, D., Ness, W. & Drenckhahn, D. *J. Cell Sci.* **110**, 765–770 (1997).
4. Sobkowicz, H. M., August, B. K. & Slapnick, S. M. *Int. J. Dev.*

- Neurosci.* **15**, 463–485 (1997).
 5. He, D. Z. Z., Zheng, J. & Dallos, P. *Hearing Res.* **162**, 113–125 (2001).
 6. Zine, A. & Romand, R. *Brain Res.* **721**, 49–58 (1996).
 7. Kaltenbach, J. A., Falzarano, P. R. & Simpson, T. R. *J. Comp. Neurol.* **350**, 187–198 (1994).

8. Nguyen-Legros, J. & Hicks, D. *Int. Rev. Cytol.* **196**, 245–313 (2000).
 9. Miller, J. D., Watson, C. S. & Covell, W. P. *Acta Otolaryngol.* **176** (suppl.), 1–89 (1963).
 10. Dumont, R. A. et al. *J. Neurosci.* **21**, 5066–5078 (2001).
Competing financial interests: declared none.

Artificial intelligence

Fast hands-free writing by gaze direction

Here we describe a method for text entry based on inverse arithmetic coding that relies on gaze direction alone and which is faster and more accurate than using an on-screen keyboard. These benefits are derived from two innovations: the writing task is matched to the capabilities of the eye, and a language model is used to make predictable words and phrases easier to write.

For people who cannot use a standard keyboard or mouse, the direction of gaze is one of the few means by which they can convey information to a computer. Many systems for gaze-controlled text entry provide an on-screen keyboard with buttons that can be ‘pressed’ by staring at them. But eyes did not evolve to push buttons, and this method of writing is exhausting.

Moreover, on-screen keyboards are inefficient because typical text has considerable redundancy¹. Although a partial solution to this defect is to include word-completion buttons as alternative buttons alongside

the keyboard, a language model’s predictions can be better integrated into the writing process. By inverting an efficient method for text compression — arithmetic coding² — we have created an efficient method for text entry, which is also well matched to the eye’s natural talent for search and navigation.

One way to write a piece of text is to delve into a theoretical ‘library’ that contains all possible books, and find the book that contains exactly the desired piece of text³; writing thus becomes a navigational task. In our idealized library, the ‘books’ are arranged alphabetically on one enormous shelf. As soon as the user looks at a part of the shelf, the view zooms in continuously on the point of gaze. So, to write a message that begins ‘hello’, the user first steers towards the section of the shelf marked ‘h’, where all the books beginning with ‘h’ are found. Within this section are different sections for books beginning ‘ha’, ‘hb’, ‘hc’ and so on; the user enters the ‘he’ section, then the ‘hel’ section within it, and so forth.

To make the writing process efficient, we use a language model, which predicts the probability of each letter’s occurrence in a given context, to allocate the shelf space

for each letter of the alphabet (Fig. 1a). When the language model’s predictions are accurate, many successive characters can be selected by a single gesture.

We previously evaluated this system, which we call ‘Dasher’, with a mouse as the steering device⁴. Novices rapidly learned to write and an expert could write at 34 words per minute; all users made fewer errors than when they were using a standard ‘QWERTY’ keyboard.

Figure 1b shows an evaluation of Dasher driven by an eye-tracker, compared with an on-screen keyboard. After an hour of practice, Dasher users could write at up to 25 words per minute, whereas on-screen keyboard users could manage only 15 words per minute. Moreover, the error rate with the on-screen keyboard was about five times that obtained with Dasher.

Users of both systems reported that the on-screen keyboard was more stressful to use than Dasher for two reasons. First, they often felt uncertain whether an error had been made in the current word (the word-completion feature works only if no error has been made); an error can be spotted only by looking away from the keyboard. Second, a decision has to be made after ‘pressing’ each character on whether to use word completion or to continue typing — looking to the word-completion area is a gamble as it is not guaranteed that the required word will be there, and finding the correct completion requires a switch to a new mental activity. By contrast, Dasher users can see simultaneously the last few characters they have written and the most probable options for the next few. Furthermore, Dasher makes no distinction between word completion and ordinary writing.

Dasher works in most languages — the language model can be trained on sample documents and adapts to the user’s language as he or she writes. It can also be operated with other pointing devices, such as a touch screen or rollerball. Dasher is potentially an efficient, accurate and fun writing system not only for disabled computer users but also for users of mobile computers.

David J. Ward, David J. C. MacKay

Cavendish Laboratory, Cambridge CB3 0HE, UK
 e-mail: mackay@mrao.cam.ac.uk

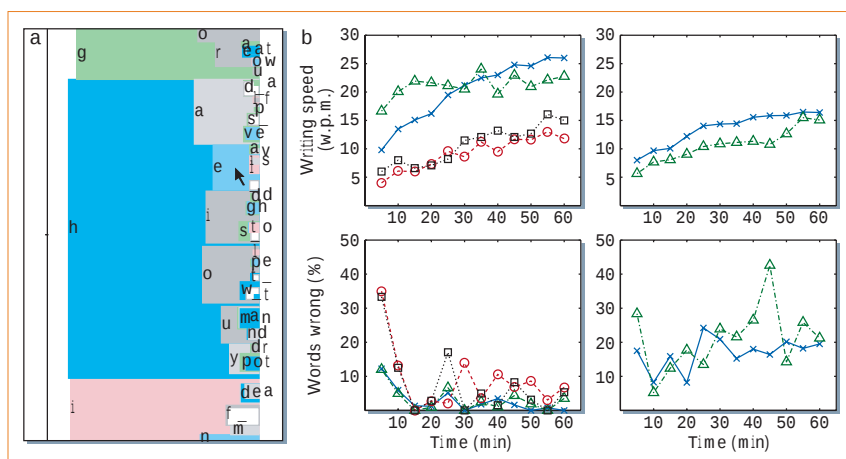


Figure 1 Hands-free text entry. **a**, Screenshot of ‘Dasher’⁵ when the user begins writing ‘hello’. The shelf of the alphabetical ‘library’ is displayed vertically. The space character (represented as an underscore) is included in the alphabet after ‘z’. In this example, the user has zoomed in on the portion of the shelf containing messages beginning with ‘g’, ‘h’ and ‘i’. Following the letter ‘h’, the language model makes the vowels and ‘y’ easier to write by giving them more space. Common words such as ‘had’ and ‘have’ are visible. The arrow indicates the gaze of the user; its vertical coordinate controls the zooming-in point and its horizontal coordinate controls the rate of zooming; looking to the left makes the view zoom out, allowing recent errors to be corrected. **b**, Comparison of writing speeds and error rates for two methods of gaze-driven text entry. Left, Dasher with eye-tracker, as recorded for two expert users of the system (crosses, triangles) and two novices (circles, squares); right, on-screen keyboard, used by two experts on the ‘QWERTY’ keyboard. The eye-tracking system was EyeTech’s Quick Glance eye-tracker. Each user took dictation from Jane Austen’s *Emma* in 5-min sessions. The language model PPMD5 predicts the next character when given the previous five characters^{6,7}; it was trained on passages from *Emma* not included in the dictation. Right panels, the two experts took dictation using the same eye-tracker to control the WvWk on-screen keyboard (a standard ‘QWERTY’ keyboard) with the word-completion buttons enabled.

- Shannon, C. E. *Bell Syst. Tech. J.* **27**, 379–423, 623–656 (1948).
- Witten, I. H., Neal, R. M. & Cleary, J. G. *Commun. Assoc. Comput. Machin.* **30**, 520–540 (1987).
- Borges, J. L. *The Library of Babel* (Godine, Boston, 1941).
- Ward, D. J., Blackwell, A. F. & MacKay, D. J. C. *Proc. 13th Annu. ACM Symp. User Interface Software Technol.* **2000**, 129–137 (2000).
- Ward, D. J. Dasher, version 1.6.4 (2001); www.inference.phy.cam.ac.uk/dasher
- Cleary, J. G. & Witten, I. H. *IEEE Trans. Commun.* **32**, 396–402 (1984).
- Teahan, W. J. *Proc. NZ Comp. Sci. Res. Stud. Conf.* (1995); citeseer.nj.nec.com/teahan95probability.html

Competing financial interests: declared none.

Scale dependence of bubble creation mechanisms in breaking waves

Grant B. Deane & M. Dale Stokes

Marine Physical Laboratory, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0238, USA

Breaking ocean waves entrain air bubbles that enhance air–sea gas flux, produce aerosols, generate ambient noise and scavenge biological surfactants. The size distribution of the entrained bubbles is the most important factor in controlling these processes, but little is known about bubble properties and formation mechanisms inside whitecaps. We have measured bubble size distributions inside breaking waves in the laboratory and in the open ocean, and provide a quantitative description of bubble formation mechanisms in the laboratory. We find two distinct mechanisms controlling the size distribution, depending on bubble size. For bubbles larger than about 1 mm, turbulent fragmentation determines bubble size distribution, resulting in a bubble density proportional to the bubble radius to the power of $-10/3$. Smaller bubbles are created by jet and drop impact on the wave face, with a $-3/2$ power-law scaling. The length scale separating these processes is the scale where turbulent fragmentation ceases, also known as the Hinze scale. Our results will have important implications for the study of air–sea gas transfer.

Breaking waves in the open ocean create dense plumes of bubbles that can extend 0.5 m or more below the surface and have void fractions of air exceeding 10%. The bubbles in the plumes range in size from tens of micrometres to centimetres, and are known to be important for a number of diverse phenomena, including the enhanced transport of gases across the ocean surface, ambient noise generation, the production of aerosols, and the scavenging of biological surfactants^{1–4}. The single most important property of the bubbles is their size distribution. The distributions and dynamics of bubbles in the upper ocean tens to hundreds of seconds after their formation are fairly well understood^{5–11}. But few field measurements have been made of bubbles during the first seconds of wave breaking^{12–15} and within the high-void-fraction region of the whitecap, and bubble formation mechanisms have not been quantified. One of the problems with existing measurements of oceanic bubble size distributions is that the bubble spectrum evolves quite rapidly in the first second or so after the active phase of air entrainment ceases, making it difficult to reconcile different data sets without knowing the exact age of the bubble plume. This rapid evolution has been described by Monahan⁸, who introduced the idea of dense ‘alpha’ and more diffuse ‘beta’ plumes.

Models of breaking-wave bubble spectra have been proposed. Longuet-Higgins¹⁶ considered the bubble size spectrum created by the random cleaving of a cube, and showed that, as the number of cleaving events becomes large, the resulting bubble distributions resemble that of observations. However, this theory does not offer a way of relating the size distribution to the physical properties of the wave. Baldy¹⁷ discussed bubble formation under the assumption that the bubble creation rate depends only on ε , the fluid turbulent dissipation rate, and the bubble formation energy, resulting in a -2 power-law scaling with radius. Garrett *et al.*¹⁸ have proposed a cascade of bubble fragmentation events with the bubble spectrum depending only on ε and the average rate of supply of air, resulting in a $-10/3$ power-law scaling. In addition, there has been significant work on bubble splitting in sheared and turbulent flows and air entrainment by plunging jets and drops (see, for example, refs 19–21).

However, there are no prior observations of bubble formation processes inside breaking waves. Photographic studies of air entrainment in laboratory waves have been reported previously^{22–24}, but at insufficient temporal and spatial scales to resolve bubble formation mechanisms within the breaking crest. The problem is making measurements on the appropriate length and time scales in the

complex two-phase flow inside a breaking wave. Using optical and acoustical observations of bubble formation, we have quantified various aspects of the phenomena determining the bubble size spectrum in laboratory and oceanic waves.

The origin of the bubble size spectrum

The lifetime of wave-generated bubbles falls into two phases. The first phase occurs as bubbles are entrained and fragmented inside the breaking wave crest. Newly created bubbles produce pulses of sound, so this period is accompanied by a burst of noise and we refer to it as the acoustically active phase. Once bubble creation processes cease, the newly formed bubble plume becomes acoustically quiescent and evolves under the influence of turbulent diffusion, advection, buoyant degassing and dissolution. In Monahan’s⁸ nomenclature, both these phases fall within the lifetime of an

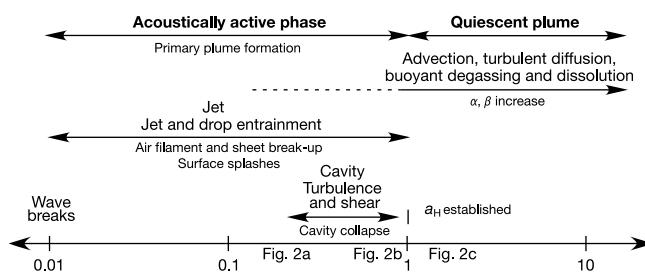


Figure 1 Logarithmic timeline of bubble plume evolution. Plume lifetime can be divided into two main phases—the acoustic phase when bubbles are formed, and the quiescent phase which begins when active bubble formation ceases. The physical processes occurring during the acoustic phase determine the initial bubble size distribution. The quiescent plume evolves rapidly under the influence of buoyant degassing, turbulent diffusion, advection and dissolution. Most oceanic bubble size distributions have been measured in quiescent plumes which have evolved from the initial size distribution. Two primary mechanisms determine the bubble size distribution during the acoustic phase. The first is jet and drop entrainment, which is active during the entire acoustic phase and determines the slope (α) of the size distribution for bubbles smaller than the Hinze scale (a_H). The second is bubble fragmentation in turbulent and sheared flow. This mechanism operates during the wave cavity collapse, and determines the slope (β) of the distribution for bubbles larger than the Hinze scale. The Hinze scale is determined by the turbulent dissipation rate within the breaking wave crest. Labels on the time axis show when the images of plume formation were taken (Fig. 2).

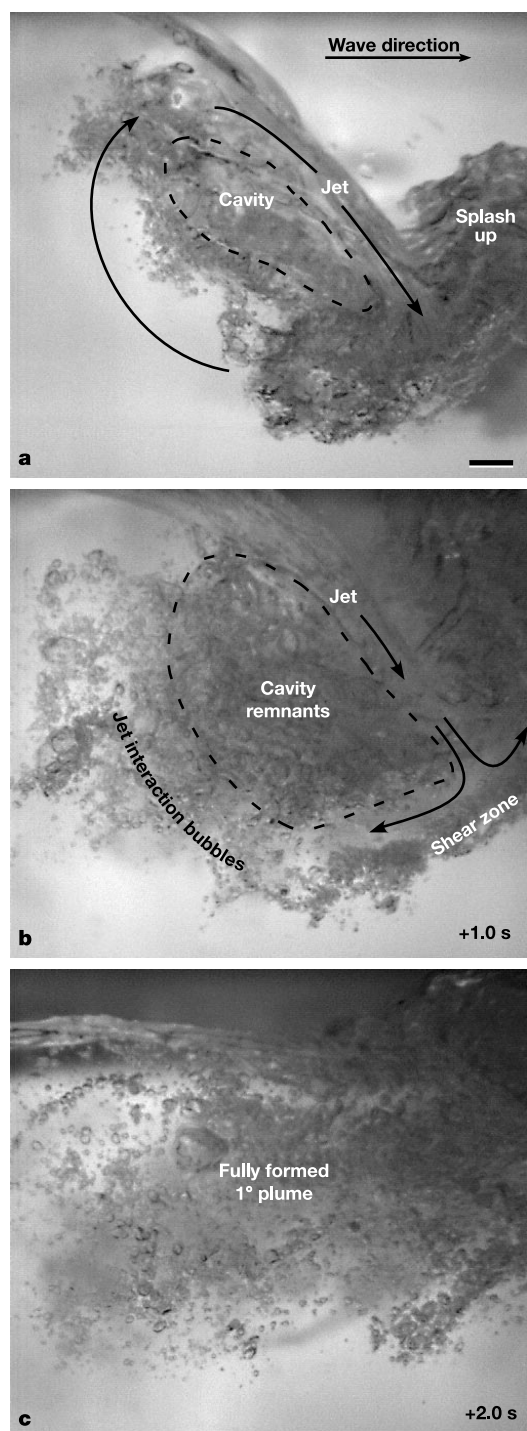


Figure 2 Three high-speed video images of a breaking wave crest taken during the acoustically active phase of the wave crest. Scale bar in **a**, 1 cm. The air–water boundaries scatter light out of the plane of illumination and appear darker. **a**, Bubble formation before the cavity of air trapped between the overturning jet and the wave face has fragmented. The bubbles around the lower half of the cavity were formed by jet and drop impact at the wave face and then advected around the cavity by clockwise flow. **b**, Bubble formation during the collapse of the air cavity. The jet-induced bubbles encircle the cavity remnants, and the jet/wave-face interaction region has developed into a shear zone. Detailed image analysis shows that bubble fragmentation occurs both inside the cavity region and in the shear zone. **c**, Bubble plume at the end of the acoustic phase. Image sequences were captured at 250 or 1,000 frames s^{-1} and a shutter speed of $1/3,000$ s using a Kodak Motioncorder SR-1000 camera.

alpha plume. Here we describe the physical processes that control bubble formation during the active phase, and determine the bubble spectrum at the transition between the acoustically active and quiescent plumes. These phases are delineated in Fig. 1, which shows a logarithmic timeline annotated with the various processes dominating bubble evolution.

We conducted a series of experiments in a seawater wave flume (33 m long, 0.5 m wide, 0.6 m water depth). Wave packets were generated at one end of the flume using a computer-controlled wave paddle, which produced plunging breakers approximately 10 cm in height. The amplitude and phase of the wave-packet frequency components were generated so that they added constructively at the wave break point, producing a breaking wave. The wave-packet centre frequency, wavelength, relative bandwidth and slope were respectively 0.73 Hz, 2.3 m, 1 and 0.4. Surface elevation time series computed from pressure measurements made upstream and downstream of the breaking region were used to compute the wave-packet energy lost owing to breaking, and to ensure that events were repeatable. Events were imaged a few centimetres (2.5–18 cm) away from the glass-walled side of the flume using a high-speed video camera and front and back lighting. Simultaneous measurements of the noise generated by the wave crest were made with a hydrophone beneath the breaking region.

The breaking wave crest produces a complicated two-phase flow that evolves over a range of length and time scales. Those flow features that form distinct and repeatable patterns are shown in Fig. 2. Figure 2a shows the flow after the overturning wave crest (plunging jet) has struck the wave face, but before the cavity of air trapped between the jet and wave face (cavity) has fragmented. The arrows indicate the clockwise circulation around the cavity caused by the wave motion. The impact of the jet on the wave face causes a reactionary splash-up, which results in the formation of a secondary bubble plume. Figure 2b shows the same wave crest 1 s later. The cavity has almost completely fragmented, and the plunging jet has formed a shear layer on the wave face. Sometimes a layer of air trapped between the spreading jet and wave face is observed to form and split into filaments and bubbles. The cavity remnants are encircled by a cloud of bubbles formed by the interaction of the plunging jet with the wave face. The clockwise circulation around the cavity advects bubbles entrained by the jet and, ultimately, some of them are re-circulated through the jet. Figure 2c shows the wave crest 2 s later. The plunging jet has collapsed, leaving a fully formed, quiescent plume of bubbles.

The images in Fig. 2 suggest two distinct flow features driving bubble creation: the jet/wave-face interaction and the collapsing cavity. This view is supported by an analysis of the underwater noise radiated by the breaking wave. Newly created bubbles act like damped resonators and emit a pulse of sound^{25,26}. The centre frequency of the sound pulse varies inversely with bubble radius (for example, a 3.3-mm-radius bubble resonates at approximately 1 kHz near the sea surface), so the breaking-wave noise is characteristic of the sizes of bubbles being created within the crest and persists while bubbles are being formed. Figure 3 shows a spectrogram of wave noise averaged over 17 breaking events and plotted as a colour contour map versus frequency and time. The acoustic frequency (F) expressed in terms of a resonant bubble radius (a) are shown on the left-hand side of the figure. The flow structures characteristic of different phases of noise emission are shown in images at the bottom of the figure.

Two distinct periods of sound production can be seen in the spectrogram. The period labelled 'Jet' is associated with the creation of bubbles from 2 mm down to at least 0.1 mm radius, and persists through the entire acoustically active phase. The shorter period, labelled 'Cavity', contains a burst of low-frequency noise centred around 300 Hz, and is associated with the creation of much larger bubbles (2 mm to ≥ 10 mm radius). Examination of video images shows that this noise is simultaneous with the breakup of the cavity.

The low-frequency sound generated by the cavity breakup is not radiated before its collapse, implying that bubbles larger than about 2 mm are produced only during cavity collapse. This is confirmed by manual, optical bubble counts taken before and during cavity collapse (not shown).

One of the primary objectives of this study was to measure the bubble size distribution within the interior of actively breaking wave crests. This has been accomplished by manually identifying and sizing bubbles in images similar to those shown in Fig. 2. Because the bubble size spectra are calculated from two-dimensional images of a three-dimensional volume, the size distribution estimates can be biased by bubble occlusion and bisection of large bubbles by the image plane. However, when the thickness of the sample section is small relative to the total area of the image, the biases are negligible²⁷.

Figure 4 shows the result of the image analysis. The bubble size distribution was estimated from approximately 225 images taken during the acoustically active period of 14 wave-breaking events. The bubble spectrum shows two distinct scaling laws (with power-law exponents α and β respectively) with a break point at slightly less than 1 mm. The power-law scaling of bubble density on radius is $-3/2$ for bubbles smaller than about 1 mm and $-10/3$ for larger bubbles. Figure 4 inset shows the temporal evolution of bubbles observed in a single plume. The upper curve was measured at the transition between the active and quiescent phases, and it shows the same power-law scaling as the ensemble average. The lower curve

was measured 1.5 s later, and both α and β show a significant increase due to bubble degassing and dissolution. This illustrates the rapid evolution of the bubble spectrum once active bubble creation processes have stopped, and provides a context for interpreting oceanic bubble spectra where the age of the plume is often uncertain.

Bubble breakup in turbulent and sheared flow

The density of bubbles larger than about 1 mm follows a $-10/3$ power-law scaling with radius. The acoustic data and visual observation of cleaving bubbles suggest that these bubbles are created as the air cavity fragments²⁴ and bubbles split in the shear zone. Garrett *et al.*¹⁸ envisage a physical process where air is entrained into relatively large bubbles, which fragment into smaller bubbles at a rate which depends on ε . Assuming that the bubble size spectrum $N(a)$ (number of bubbles per m^3 per μm radius increment) depends only on the average rate of supply of air Q (the volume of air entrained per volume of water per second) with dimensions s^{-1} , ε and radius a , then dimensional consistency implies the scaling law

$$N(a) \propto Q \varepsilon^{-1/3} a^{-10/3} \quad (1)$$

which is consistent with the observed scaling law for bubbles larger than about 1 mm.

The central idea behind turbulent fragmentation^{28,29} is that a bubble is likely to break up if the differential pressure forces across the bubble exceed the restoring forces of surface tension. The ratio of these forces is the Weber number, given by

$$We = (\rho/\gamma)u^2d \quad (2)$$

where³⁰ ρ is the fluid density, γ is the fluid surface tension, u is the turbulent velocity field on the scale of the bubble and d is the bubble diameter. The condition for bubble fragmentation is that We exceed a critical value, We_c . Recent experiments suggest that We_c lies in the range 3–4.7 (refs 30 and 31), and we have used the value $We_c = 4.7$. An additional constraint is that the bubble undergo shape oscill-

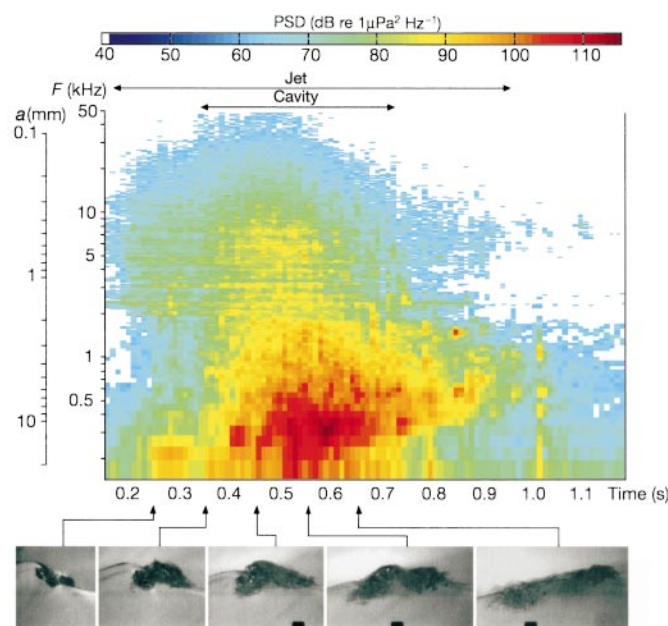


Figure 3 Spectrogram of wave noise calculated from an average of 17 breaking events. This spectrogram is similar to those recorded³⁷ during the collective oscillations of bubble plumes. The colour contours represent sound intensity plotted on a decibel scale (the intensity is referenced to $1 \mu\text{Pa}^2 \text{Hz}^{-1}$) versus frequency and time. The log scale labelled 'a' on the left-hand side indicates the radius of a bubble resonant at the corresponding frequency (F) on the frequency scale. The two time periods labelled 'Jet' and 'Cavity' are discussed in the text. The wave noise was measured with a hydrophone (International Transducer Corp. 6050c) mounted in the wave flume beneath the bubble plumes. Because the flume is a bounded enclosure, it exhibits acoustic resonances that impose structure on the wave noise spectrum. The amplitude, centre frequencies and spectral widths of the resonant bands were estimated by placing a broad-band acoustic source at the primary plume location and measuring the tank response. These estimates were then used to compensate for the effect of the acoustic resonances on the breaking-wave noise spectrum. The images plotted beneath the spectrogram show the sequence of flow features observed at different times during the acoustic emission. The black bar is 40 mm long.

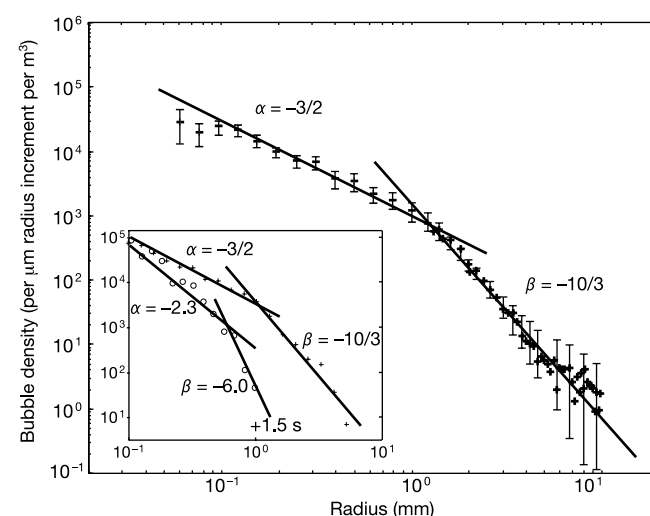


Figure 4 The average bubble size spectrum estimated from 14 breaking events during their acoustic phase. Two camera magnifications were used and the results superimposed to obtain the slightly greater than two decades of bubble radii observed. The vertical scale is number of bubbles per m^3 in a bin radius $1 \mu\text{m}$ wide. Vertical bars show ± 1 s.d. The size distribution shows a marked change in slope at a radius that we are identifying as the Hinze scale. Bubbles larger and smaller than this scale respectively vary as $(\text{radius})^{-10/3}$ and $(\text{radius})^{-3/2}$ denoted by β and α . Inset, the bubble size distribution at the beginning of the quiescent phase (crosses) and 1.5 s into the quiescent phase (open circles). Both slopes of the bubble spectrum have increased noticeably during this time interval. This rapid evolution becomes important when interpreting size distributions collected during the plume quiescent phase.

lations³² so that the ratio of the bubble major-to-minor axes (the bubble eccentricity) exceeds ~ 3 . This constraint is satisfied if the bubble Reynolds number, given by

$$Re = ud/\xi \quad (3)$$

where ξ , the fluid kinematic viscosity, is greater than a critical value³² (~ 450). If the fluctuating velocity field is assumed to be described by Kolmogorov's inertial subrange, so that $u^2 = 2\varepsilon^{2/3}d^{2/3}$,

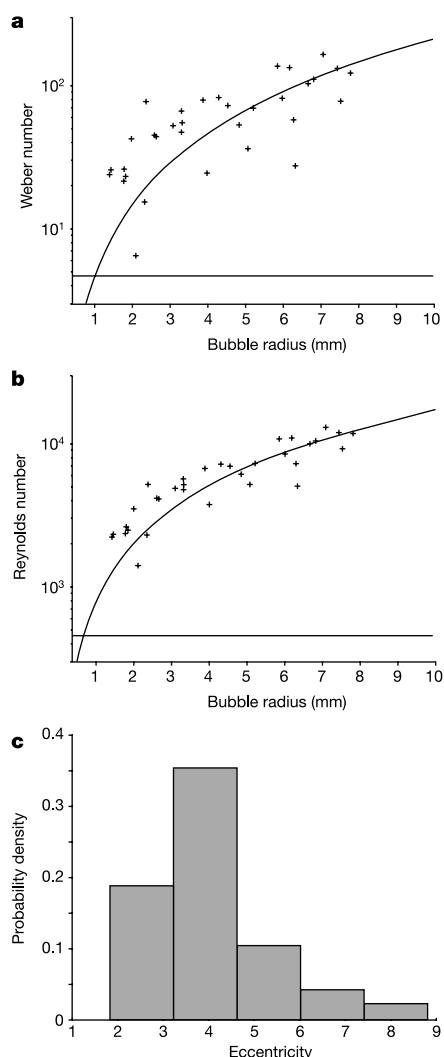


Figure 5 Some bubble fragmentation metrics. Fragmenting bubbles were identified and analysed in image sequences of breaking events. The velocity shear across a splitting bubble was estimated from the separation velocity of the fragmentation products. **a**, The observed Weber number as a function of parent bubble radius. The horizontal line indicates the critical Weber number (4.7) that must be exceeded for fragmentation to occur. The curve shows the theoretical Weber number expected for fully developed turbulence with an energy dissipation rate $\varepsilon = 12 \text{ W kg}^{-1}$. **b**, The observed Reynolds number for flow across the parent bubble. The horizontal line shows the critical Reynolds number (~ 450) that must be exceeded for bubble shape oscillations and elongation, and subsequent fragmentation to occur. The curve shows the expected Reynolds number for the same dissipation rate as above. **c**, The probability density of bubble eccentricity (the ratio of semimajor to semiminor axis) measured just before bubble separation. The mean eccentricity is about 4. Inset, images illustrating the fragmentation of a 4.5-mm-radius bubble in the shear zone of the jet/wave-face interaction region. The arrows in the top two images indicate the fragmentation point, and the two arrows in the bottom image show the fragmentation products.

then only bubbles with radius larger than the Hinze scale

$$a_H = 2^{-8/5} \varepsilon^{-2/5} (\gamma W e_c / \rho)^{3/5} \quad (4)$$

will fragment. The Reynolds number also implies a critical length scale, but the values of γ and ξ for sea water are such that it is smaller than a_H for reasonable values of ε . These two scales become comparable only when $\varepsilon \approx 330 \text{ W kg}^{-1}$.

In order to test these ideas, we examined sequences of images from breaking waves to identify fragmenting bubbles. Although fragmentation events were frequently observed in the collapsing cavity and shear zone, image sequences of fragmenting bubbles suitable for quantitative analysis were relatively rare: 33 usable events were found in about 25,000 video frames from 125 breaking-wave events. Fragmentation metrics were estimated for each event. The differential velocity across a splitting bubble was estimated by measuring the relative velocity of separation of the fragmentation products, and the bubble eccentricity was measured just before fragmentation. The measured separation velocity was always a factor of 5 or more greater than the expected rise velocity of the bubble products, and buoyancy effects were therefore assumed to be an unimportant source of bias. Figures 5a and 5b show scatter plots of the measured Weber and bubble Reynolds numbers as a function of bubble radius. All the observed fragmenting bubbles were larger than 1 mm in radius, and the estimated Weber and Reynolds numbers were greater than the critical values described above (shown as horizontal lines on the lower portion of the plots). The histogram in Fig. 5c shows the observed bubble eccentricity probability density, and the three images on the right-hand side illustrate the elongation that occurs as a bubble fragments. The observed eccentricity is consistent with the idea that bubbles must reach a critical degree of distortion before they fragment³¹.

The turbulent energy dissipation rate inside the fragmenting cavity region can be estimated from equation (4) if the Hinze scale is known. Because bubbles smaller than the Hinze scale are stabilized by surface tension forces, a reasonable estimate for this scale is the radius of the smallest bubbles observed to fragment. This scale estimate differs from that adopted by Garrett *et al.*¹⁸, who used the definition³³ that 95% of the air is contained in bubbles with a radius less than a_H . This definition may be more appropriate for steady-state entrainment (for example, steady-state jets), where bubbles have sufficient time to completely fragment. In the present case, these target bubbles are in the process of fragmenting.

The smallest bubbles we observed fragmenting were about 1 mm in radius. An estimate of 1 mm for the Hinze scale is also consistent with the observed change in power-law scaling of the bubble spectrum at around 1 mm (Fig. 4). Taking this to be the Hinze scale, equation (4) yields a value of $\varepsilon = 12 \text{ W kg}^{-1}$. The expected Weber and bubble Reynolds numbers as a function of bubble radius for this value of ε calculated from equations (2) and (3) are plotted as solid curves in Fig. 5a and b, and are in good agreement with the observed values. We can also estimate ε from the results of ref. 34, where the normalized dissipation rate per unit length of crest in unsteady breaking waves (Θ) was measured: $\Theta = \varepsilon_l g / \rho_w C^5$, where ε_l is the dissipation rate per unit length of crest, g is the acceleration due to gravity, ρ_w is the fluid density, and C is the wave speed. These authors measured a value of $\Theta \approx 0.009$ for waves with multiple breaks (like ours). Our measured wave speed was 1.7 m s^{-1} , implying a dissipation rate of $\varepsilon_l = 13.5 \text{ W m}^{-1}$. Assuming that all the energy dissipation occurs inside a cylinder 2.5 cm in radius (Fig. 2b) and a void fraction of air within the collapsing region¹ of 0.5, this dissipation rate is equivalent to $\varepsilon = 13 \text{ W kg}^{-1}$, which is in good agreement with our dissipation estimate based on the Hinze scale.

Bubble entrainment

Large numbers of bubbles smaller than the Hinze scale are entrained, and here we consider their origin. Such bubbles are stabilized by surface tension forces, and do not fragment. Bubbles

comparable to the Hinze scale do fragment and produce fragmentation products smaller than the Hinze scale. However, we did not observe any fragmentation products less than $a_H/2$, suggesting they are not a major source of small bubbles. The wave noise spectrogram (Fig. 3) shows that bubbles from around 2 mm radius to bubbles at least as small as 100 μm are created at the start of the active phase, before the cavity fragments, and persist to the end of this phase. The main flow feature associated with these bubbles is the interaction of the plunging jet (including the impact of drops) formed by the overturning wave crest with the wave face (Fig. 2a). The implication is that most bubbles smaller than the Hinze scale are formed by the jet/wave-face interaction throughout the active phase.

There is a significant body of literature describing air entrainment by steady-state, laminar and turbulent plunging jets^{19–21}. Although jet entrainment in breaking waves is by transient, inclined jets moving relative to the water surface, it is possible to derive a power scaling law for the bubble spectrum from the mechanics of steady-state jets. Following dimensional arguments similar to those used by Garrett *et al.*¹⁸ to derive their scaling law for turbulent fragmentation, we will assume that the bubble spectral level is a multiplicative function of γ , ρ , a and jet velocity v . Surface tension (γ) is included because we are considering bubbles smaller than the Hinze scale where surface tension effects are assumed to be important. The jet velocity is included to account for the fact that low-velocity jet air entrainment rates are known to scale¹⁹ with v^2 . For dimensional consistency, the bubble size spectrum must then scale as:

$$N(a) \propto Q(\gamma/\rho)^{-3/2} v^2 a^{-3/2} \quad (5)$$

Although we do not have a mechanistic argument supporting equation (5), the $-3/2$ power scaling law for bubble spectral density with bubble radius is in good agreement with the observed spectral slope for bubbles smaller than the a_H . A greater understanding of the deterministic entrainment physics would require the kind of detailed analysis presented by, for example, ref. 21 for bubble entrainment by single drop impacts.

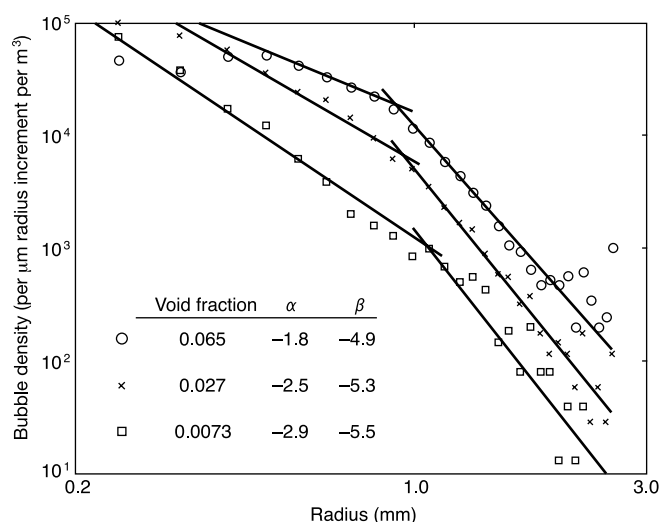


Figure 6 Oceanic bubble size distributions observed 30 cm below whitecaps during the plume quiescent phase. Each curve was obtained by analysing multiple images from a single breaking event. The exact time after the beginning of the quiescent phase is unknown, but is of the order of a second. Using the void fraction of air as a surrogate for plume age, the size distribution slopes α and β can be seen to increase with increasing plume age as bubbles are lost to the imaging volume through the effects of buoyant degassing and dissolution.

Oceanic waves

The main interest in bubble size spectra beneath breaking waves is their relevance to oceanic phenomena, such as air–sea gas exchange. It is important, then, to consider the relationship between our laboratory results and air entrainment processes occurring in open-ocean waves where different routes to breaking exist, and different wave amplitudes and frequencies are encountered.

Ideally, an analysis of bubble formation mechanisms in oceanic waves would be based on visual images of the flow features occurring inside whitecaps and measurements of bubble fragmentation and entrainment. However, the best observations of open-ocean whitecaps available at present are oceanic bubble spectra from spilling breakers collected a second or so into the quiescent plume phase approximately 30 cm beneath the surface¹⁴. Despite this limitation, we can make some inferences by comparing the oceanic bubble spectra with those observed in the laboratory.

Figure 6 shows typical bubble size spectra from three separate breaking events observed off the coast of Southern California during the winter of 2000. The wind speed, significant wave height, wave period and water temperature were respectively 7–10 m s^{-1} , 2.5–3.5 m, 9–12 s and 13.2 °C. The bubble spectra from the interiors of whitecap plumes were observed using an optical bubble counting instrument²⁶ mounted on a surface-tracking frame deployed from the research platform FLIP. It was impossible to determine the exact end of the active phase during the breaking events and so the quiescent plume age cannot be determined precisely, but is of the order of 1 s. Underwater video footage of oceanic whitecaps and our flume study show that the plume has well-defined spatial boundaries at the end of the active phase. As a first approximation, the effect of spatial heterogeneity on void fraction can be neglected, and relative plume age can be determined using void fraction of air as a surrogate for time into the quiescent phase¹. Under this assumption, plume age increases from the top to the bottom of the figure.

The most revealing aspect of the oceanic spectra is that, like the flume spectra, they exhibit two power-law scales and an increase in slope at around 1 mm bubble radius. This change in slope has also been observed in quiescent plume size distributions measured beneath breaking surf³⁵. The existence of a break-point in the oceanic spectra so similar to that observed in the laboratory suggests that the Hinze scale also exists in the open ocean where spilling breakers predominate. This oceanic Hinze scale is evidence that, as with the laboratory waves, turbulent fragmentation drives the large end of the bubble spectrum.

Although the open-ocean whitecaps and breaking surf were different from the laboratory waves in terms of amplitude, frequency and breaker type, they have similar Hinze scales. There are reasons to expect that this would be the case. Measurements³⁴ show that dissipation rates remain relatively constant for multiple break events over a range of integral wave slopes and, from equation (4), the Hinze scale shows a relatively weak dependence on the turbulent dissipation rate. A change in dissipation rate from 3 to 30 W kg^{-1} corresponds to a change in Hinze scale from 1.7 to 0.7 mm.

The second revealing aspect of the oceanic spectra is their power-law scaling. We would expect to observe $\alpha = -3/2$ and $\beta = -10/3$ for bubble spectra measured at the end of the active phase, and slopes greater than these for older plumes (Fig. 4 inset). In fact, we do observe similar, but somewhat greater slopes, and plumes with increasing age (inferred using their void fraction) show increasing slopes with increasing time.

The importance of the Hinze scale

We have presented qualitative and quantitative experimental evidence that two primary mechanisms control the numbers and sizes of bubbles created by laboratory and oceanic breaking waves. The physical mechanisms operate on different regions of the bubble size spectrum, separated by the Hinze scale. Bubbles larger than the

Hinze scale are subject to fragmentation by turbulent and sheared flow, and show a $-10/3$ power-law scaling with radius. Bubbles smaller than the Hinze scale are stabilized by surface tension, and show a $-3/2$ power-law scaling with radius. The $-10/3$ power law scaling is consistent with recent theoretical predictions made by Garrett *et al.*¹⁸, and the $-3/2$ power law scaling is consistent with a dimensional analysis of jet entrainment presented here.

Despite the fact that breaking in the open ocean is predominately spilling, rather than plunging, a comparison of open-ocean spectra with those observed in the laboratory show important similarities. There is clear evidence of a Hinze scale in the oceanic bubble spectra, and spectral slopes are consistent with the evolving slopes of quiescent plumes. The fact that we see the Hinze scale in both open-ocean and surf-zone spectra, arising from spilling and plunging breakers with very different scales, implies that the same bubble formation mechanisms are operating in each of these environments. This conclusion is supported by the spectral slopes observed in quiescent oceanic plumes, which are similar to, but somewhat greater than, the slopes observed at the end of the active phase in the laboratory. This increase in slope with increasing time is expected, as bubbles are lost through buoyant degassing and become smaller as they dissolve. We also note that Loewen *et al.*³⁶ have measured bubble spectra immediately behind the crests of gently spilling laboratory waves. The slope of their distribution for bubbles larger than 1 mm are similar to our flume and open-ocean results, supporting the argument that the mechanism that creates large bubbles is the same beneath spilling and plunging breakers.

The existence of simple scaling laws for bubble number density separated by a length scale that depends only on the turbulent dissipation rate has important implications for modelling turbulent, two-phase flow in a wide variety of natural and man-made systems, including modelling the bubble-mediated transfer of greenhouse gases and aerosol production, both important for global climate change, and models of wave-induced oceanic ambient noise. □

Received 15 May; accepted 4 July 2002; doi:10.1038/nature00967.

1. Lamarre, E. & Melville, W. K. Air entrainment and dissipation in breaking waves. *Nature* **351**, 469–472 (1991).
2. Wallace, D. W. R. & Wirick, C. D. Large air–sea fluxes associated with breaking waves. *Nature* **356**, 694–696 (1992).
3. Farmer, D. M., McNeil, C. L. & Johnson, B. D. Evidence for the importance of bubbles in increasing air–sea gas flux. *Nature* **361**, 620–623 (1993).
4. Melville, W. K. The role of surface-wave breaking in air–sea interaction. *Annu. Rev. Fluid. Mech.* **28**, 279–321 (1996).
5. Thorpe, S. A. The effect of Langmuir circulation on the distribution of submerged bubbles caused by breaking wind waves. *J. Fluid. Mech.* **142**, 151–170 (1984).
6. Medwin, H. & Breitz, N. D. Ambient and transient bubble spectral densities in quiescent seas and under spilling breakers. *J. Geophys. Res.* **94**, 12751–12759 (1989).
7. Hwang, P. A., Hsu, Y. L. & Wu, J. Temperature effects on generation and entrainment of bubbles induced by a water jet. *J. Phys. Oceanogr.* **20**, 19–28 (1990).
8. Monahan, E. C. & Lu, M. Acoustically relevant bubble assemblages and their dependence on meteorological parameters. *IEEE J. Ocean. Eng.* **15**, 340–349 (1990).
9. Thorpe, S. A. Bubble clouds and the dynamics of the upper ocean. *Q. J. R. Meteorol. Soc.* **118**, 1–22 (1992).
10. Thorpe, S. A., Bowyer, P. & Woolf, D. K. Some factors affecting the size distributions of oceanic bubbles. *J. Phys. Oceanogr.* **22**, 382–389 (1992).

11. Farmer, D. & Li, M. Patterns of bubble clouds organized by Langmuir circulation. *J. Phys. Oceanogr.* **25**, 1426–1440 (1995).
12. Bezzabotnov, V. S., Bortkovskiy, R. S. & Timanovskiy, D. F. On the structure of the two-phase medium generated at wind-wave breaking. *Izvest. Atmos. Ocean. Phys.* **22**, 922–928 (1986).
13. Deane, G. B. & Stokes, M. D. Air entrainment processes and bubble size distributions in the surf zone. *J. Phys. Oceanogr.* **29**, 1393–1403 (1998).
14. Stokes, M. D., Deane, G. B., Vagle, S. & Farmer, D. in *Gas Transfer at Water Surfaces* Geophysical Monographs 127 (eds Donelan, M. A., Drennan, W. M., Saltzman, E. S. & Wanninkhof, R.) 278–285 (American Geophysical Union, Washington, DC, 2002).
15. de Leeuw, G. D. & Cohen, L. H. in *Gas Transfer at Water Surfaces* Geophysical Monographs 127 (eds Donelan, M. A., Drennan, W. M., Saltzman, E. S. & Wanninkhof, R.) 271–277, (American Geophysical Union, Washington, DC, 2002).
16. Longuet-Higgins, M. S. The crushing of air cavities in a liquid. *Proc. R. Soc. Lond. A* **439**, 611–626 (1992).
17. Baldy, S. A generation-dispersion model of ambient and transient bubbles in the close vicinity of breaking waves. *J. Geophys. Res.* **98**, 18277–18293 (1993).
18. Garrett, C., Li, M. & Farmer, D. The connection between bubble size spectra and energy dissipation rates in the upper ocean. *J. Phys. Oceanogr.* **30**, 2163–2171 (2000).
19. Van De Sande, E. & Smith, J. M. Jet break-up and air entrainment by low velocity turbulent water jets. *Chem. Eng. Sci.* **31**, 219–224 (1976).
20. Bonetto, F. & Lahey, R. T. Jr An experimental study on air carryunder due to a plunging jet. *Int. J. Multiphase Flow* **19**, 281–294 (1993).
21. Pumphrey, H. C. & Elmore, P. A. The entrainment of bubbles by drop impacts. *J. Fluid Mech.* **220**, 539–567 (1990).
22. Koga, M. Bubble entrainment in breaking wind waves. *Tellus* **34**, 481–489 (1982).
23. Loewen, M. R., O'Dor, M. A. & Skafel, M. G. Bubbles entrained by mechanically generated breaking waves. *J. Geophys. Res.* **101**, 20759–20820 (1996).
24. Bonmarin, P. Geometric properties of deep water breaking waves. *J. Fluid Mech.* **209**, 405–433 (1989).
25. Minnaert, M. On musical air bubbles and the sounds of running water. *Phil. Mag.* **16**, 235–248 (1933).
26. Leighton, T. G. *The Acoustic Bubble* (Academic, San Diego, 1994).
27. Stokes, M. D. & Deane, G. B. A new optical instrument for the study of breaking waves at high void fractions within breaking waves. *IEEE J. Ocean. Eng.* **24**, 300–311 (1999).
28. Kolmogorov, A. N. O drobenii kapel' v turbulentnom potoke. *Dokl. Acad. Nauk USSR* **66**, 825–828 (1949).
29. Hinze, J. O. Fundamentals of the hydrodynamic mechanism of slitting in dispersion processes. *J. Am. Inst. Chem. Eng.* **1**, 289–295 (1955).
30. Lewis, D. A. & Davidson, J. F. Bubble splitting in shear flow. *Trans. Inst. Chem. Eng.* **60**, 283–291 (1982).
31. Martínez-Bazán, C., Montañés, J. L. & Lasheras, J. C. On the breakup of an air bubble injected into a fully developed turbulent flow. I. Breakup frequency. *J. Fluid Mech.* **401**, 157–182 (1999).
32. Fan, L.-S. & Tsuchiya, K. *Bubble Wake Dynamics in Liquids and Liquid-Solid Suspensions* (Butterworth-Heinemann, Boston, 1990).
33. Clay, P. H. The mechanism of emulsion formation in turbulent flow: theoretical part and discussion. *Proc. R. Acad. Sci.* **43**, 979–990 (1940).
34. Loewen, M. R. & Melville, W. K. Microwave backscatter and acoustic radiation from breaking waves. *J. Fluid Mech.* **224**, 601–623 (1991).
35. Deane, G. B. Sound generation and air entrainment by breaking waves in the surf zone. *J. Acoust. Soc. Am.* **102**, 2671–2689 (1997).
36. Loewen, M. R., O'Dor, M. A. & Skafel, M. G. in *Third Int. Symp. on Air–Water Gas Transfer* (eds Jahne, B. & Monahan, E. C.) 337–349 (AEON Verlag & Studio, Hanau, Germany, 1995).
37. Loewen, M. R. & Melville, W. K. An experimental investigation of the collective oscillations of bubble plumes entrained by breaking waves. *J. Acoust. Soc. Am.* **95**, 1329–1343 (1992).

Acknowledgements

We thank G. Devine and J. Uyloan for assistance in laboratory data analysis, and the crew of RP *FLIP* during oceanic deployments. We also thank D. Farmer, M. Li and C. Garrett for discussions. This work was supported by the National Science Foundation and the Office of Naval Research.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.B.D. (e-mail: grant@mpl.ucsd.edu).

Directionally selective calcium signals in dendrites of starburst amacrine cells

Thomas Euler*, Peter B. Detwiler† & Winfried Denk*

* Max Planck Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Germany

† University of Washington, Department of Physiology & Biophysics, Box 357290, Seattle, Washington 98195, USA

The detection of image motion is fundamental to vision. In many species, unique classes of retinal ganglion cells selectively respond to visual stimuli that move in specific directions. It is not known which retinal cell first performs the neural computations that give rise to directional selectivity in the ganglion cell. A prominent candidate has been an interneuron called the 'starburst amacrine cell'. Using two-photon optical recordings of intracellular calcium concentration, here we find that individual dendritic branches of starburst cells act as independent computation modules. Dendritic calcium signals, but not somatic membrane voltage, are directionally selective for stimuli that move centrifugally from the cell soma. This demonstrates that direction selectivity is computed locally in dendritic branches at a stage before ganglion cells.

A substantial fraction of the neurons (retinal ganglion cells) that convey visual information out of the retina to the visual areas of the cerebral cortex respond with a strong spike discharge to a visual stimulus moving in a certain (preferred) direction but stay almost silent when the stimulus moves in the opposite (null) direction. They are therefore called 'direction-selective (DS) ganglion cells'. Even though DS ganglion cells were first described nearly 40 years ago¹ the biophysical mechanism and cellular locus for direction discrimination have been debated ever since (reviewed in ref. 2).

One of the central questions has been whether motion direction is computed in the DS ganglion cell itself (postsynaptically), using synaptic inputs that are not direction-selective, or whether the crucial computational step occurs at an 'earlier' (presynaptic) point in the signal pathway, that is, in cells that provide synaptic input to DS ganglion cells. The leading candidate for such a presynaptic locus has long been the starburst amacrine cell^{3,4}, a retinal interneuron with a highly symmetrical dendritic morphology, a peculiar distribution of inputs and outputs⁵ (Fig. 1b), and an unusually high tiling overlap factor⁶.

The role of starburst cells in direction detection became controversial when experiments using a laser to ablate starburst cells seemed to rule out their essential participation in direction selectivity⁷. Subsequently, electrophysiological evidence was presented both in favour⁸ of and against⁹ direction selectivity being computed at a presynaptic locus. Finally, recent experiments showing that direction selectivity (tested in several different ways) is abolished by genetically targeted elimination of starburst cells¹⁰ offer strong evidence that they are necessary for direction detection.

Starburst cells—like most retinal interneurons, but unlike interneurons in other parts of the central nervous system—are non-spiking. Like most amacrine cells, starburst cells possess no clear distinction between dendrite and axon, with synaptic outputs and inputs intermingled on neuronal processes (which are still called dendrites). Dendritic outputs allow locally computed information to be transferred to other cells without leaving any detectable signature in the electrical signal recorded from the soma¹¹. To study locally computed signals it is, therefore, necessary to monitor dendritic activity directly using either electrical or optical recording techniques. In this work we have examined the role of starburst-cell dendrites in detecting the direction of moving stimuli by using two-photon fluorescence measurements to monitor dendritic signals evoked by visual stimuli.

Optical recording in light-sensitive retina

Cells in the isolated intact rabbit retina were visualized with a two-photon laser scanning fluorescence microscope¹². A schematic drawing of the set-up is shown in Fig. 1a. Cell bodies were visualized using sulphorhodamine in the bath solution as a counterstain^{13,14}. We identified 'ON' starburst amacrine cells on the basis of their soma's shape (circular, diameter ~10 µm), location in the ganglion cell layer and depolarizing electrical response to the onset of illumination. A cell was selected and impaled under fluorescence-imaging control with a sharp microelectrode. Membrane potential (V_m) responses to visual stimuli were recorded and the cell was filled with fluorescent calcium indicator dye through the recording electrode, making it possible to visualize the full extent and fine structure of its neuronal morphology (Fig. 1b). The soma typically gives rise to 4–6 primary dendrites, which travel a short distance before dividing into a profusion of smaller branches with proximal, intermediate and distal zones⁵. The calcium dynamics in amacrine cells are particularly interesting because the output synapses in starburst cells are mainly restricted to the distal dendrites^{5,15}, where numerous varicosities are also apparent (Fig. 1b).

Two-photon excitation allows fluorescence imaging while leaving the retina responsive to visual stimuli because the infrared laser light (at a wavelength of about 930 nm) is very poorly absorbed by photopigments. Microfluorimetry of indicator dyes can thus be used to study light-evoked changes of the cytosolic calcium concentration ($[Ca^{2+}]$) in dendrites in the inner retina¹⁶. Figure 1c shows an example of the light-flash-triggered increase in fluorescence in a small region of a starburst-cell distal dendrite. Fluorescence intensity changes, reflecting changes in intracellular $[Ca^{2+}]$, could be evoked by visual stimuli, but also occurred spontaneously during uniform but constant illumination. Figure 1d shows examples of such spontaneous events recorded simultaneously from neighbouring dendritic branches. $[Ca^{2+}]$ transients appeared to be uncorrelated at the two optical recording sites; an early indication that dendrites can act independently of each other. The traces in Fig. 1d also demonstrate that spontaneous events can be either slow or rapid transients. The kinetics of the fast events are reminiscent of large spontaneous depolarizing potentials that are a common feature of electrical recordings from somata of starburst cells^{17–19}.

The amplitude of the $[Ca^{2+}]$ changes evoked by visual stimuli varied depending on the distance of the recording site from the soma (Fig. 1e). The changes in $[Ca^{2+}]$ were small in the proximal and intermediate zones of the dendrite and large in the distal zone. In distal portions of the dendrites light triggered either large rapid

increases in $[Ca^{2+}]$ (Fig. 1e and Fig. 2, right trace) or smaller slower decreases (see Fig. 3), depending on the stimulus location. Because synaptic vesicle exocytosis is calcium-dependent and most output synapses are located in the distal zone of starburst cell dendrites, our remaining $[Ca^{2+}]$ measurements concentrated on that region.

Localization of dendritic signals

We investigated the receptive-field structure of voltage and $[Ca^{2+}]$ responses using 'pie-wedge'-shaped sector stimuli (see Methods and the drawings at the top of Fig. 2) designed to test the independence of responses in the different dendritic branches. Although voltage responses (Fig. 2a, bottom traces) at the soma were only weakly dependent on the angular location of the stimulus, as would be expected from the symmetry of the dendritic tree, calcium increases were evoked by central sector stimuli only if the illuminated area included the dendritic branch where the recording site was located (Fig. 2). Depending on the spatial extent of the dendritic branch, one or at most two sectors elicited a $[Ca^{2+}]$ response at a particular recording site. This shows that, at least in terms of their $[Ca^{2+}]$ response, different dendritic branches can react to light stimuli locally and independently of each other, as predicted by some starburst cell models^{20,21}. The electrical responses evoked by sur-

round stimulation were inhibitory (Fig. 3a, bottom traces) and included in a few cases a delayed excitatory (OFF) response (data not shown), comparable to those described in patch-clamp voltage recordings^{19,22}. $[Ca^{2+}]$ transients were more variable between cells and between different recording sites in the same cell when stimulated with surround sectors. In most cases, illumination of the surround caused a small, slowly developing decrease in $[Ca^{2+}]$ that was followed by a delayed (about 100–150 ms) recovery (Fig. 3a) and occasionally a strong overshoot at the termination of the stimulus (OFF response; Fig. 3b). $[Ca^{2+}]$ decreases evoked by surround stimuli were not necessarily restricted to ipsilateral sectors but could be evoked at all stimulus orientations (Fig. 3b), confirming that the cells' inhibitory input is less localized than its excitatory input^{19,22}. The decrease in $[Ca^{2+}]$ was not blocked by the GABA_A-receptor antagonists bicuculline (Fig. 3c) and SR-95531 (data not shown), which implies that this decrease is mediated by a different inhibitory pathway—such as via the suppression of tonic input from bipolar cell terminals by GABA_C-receptor-mediated presynaptic inhibition^{18,23–25}.

Dendritic responses to moving stimuli

Having established that dendritic branches can act independently of

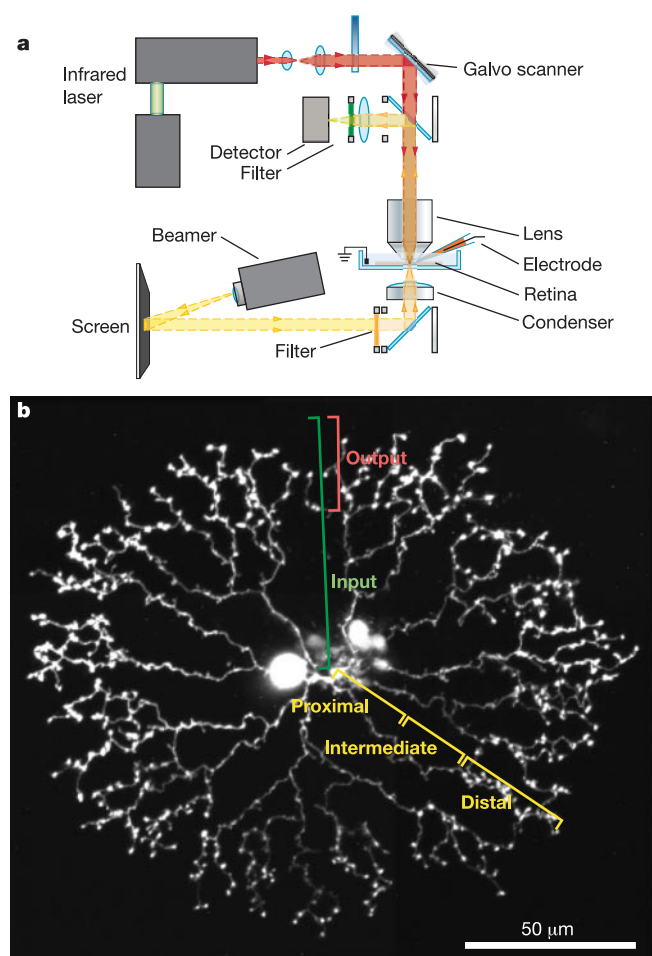
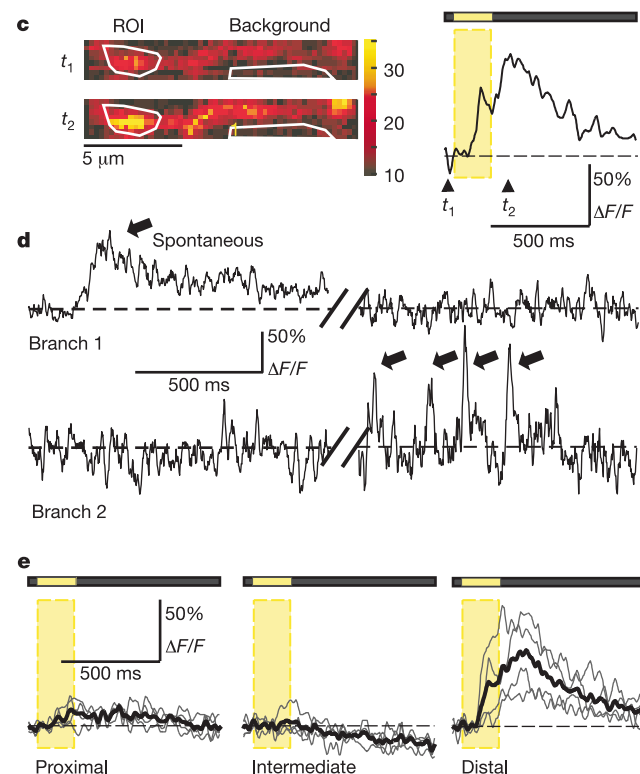


Figure 1 Calcium responses in starburst amacrine cell dendrites. **a**, Two-photon laser scanning microscope with light stimulation and electrophysiological recording capabilities. **b**, ON starburst amacrine cell in a flat-mounted rabbit retina filled with Oregon Green BAPTA-1. **c**, Dendritic $[Ca^{2+}]$ signals were monitored either with fast-frame scanning (c) or with line scans (not shown). The left panels (digitizer units) show a dendritic branch before (t_1) and just after (t_2) a light stimulus; the entire time course is



shown on the right (relative intensity change, $\Delta F/F$); for the region of interest (ROI). The background area was used for baseline correction. Vertical yellow bar and yellow section of the bar above trace indicate stimulus duration. **d**, Spontaneous calcium events (arrows) with different timing and kinetics in neighbouring branches. **e**, Calcium signals evoked by stimuli flashed over dendrite recorded at three sites (13, 110 and 233 μ m from the soma) along the same branch.

each other we investigated whether individual branches can discriminate between different directions of image motion. The somatic voltage responded to moving stimuli but did so without any obvious directional preference (Fig. 4a–c). Its main characteristic is a modulation in synchrony with the changing local light level caused by the movement of the grating. In contrast, dendritic $[Ca^{2+}]$ signals did show a strong directional dependence in response to both drifting gratings and moving bars (Fig. 4a, b, d). Grating movement in some directions triggered transient $[Ca^{2+}]$ increases (Fig. 4a) that were modulated in synchrony with the light–dark periodicity of the stimulus. Movement of the grating in the opposite directions, however, had no effect on the intracellular $[Ca^{2+}]$ at that recording location. The voltage and $[Ca^{2+}]$ responses generated by moving bright bars (Fig. 4b) were more complex. The electrical response to moving bars (top traces in each pair) began with an initial hyperpolarization as the bar entered the cell's surround (see icons at top). Next followed a depolarization as the bar moved over the cell's dendritic tree and then again a hyperpolarization as the bar

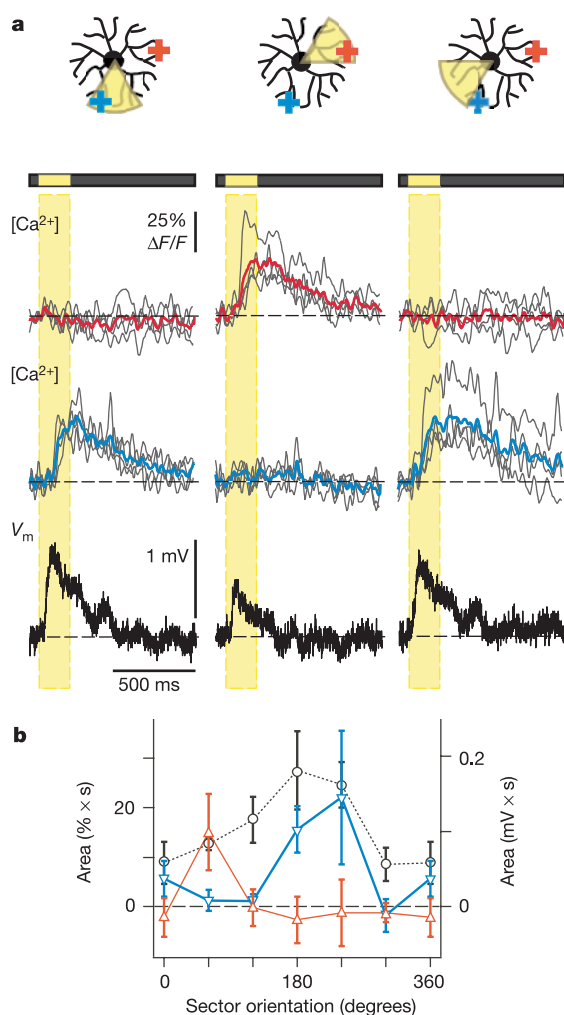


Figure 2 Local stimulation evokes local calcium signals. Somatic voltage (V_m) and dendritic $[Ca^{2+}]$ in response to 'pie-wedge'-shaped 60°-sector stimuli flashed for 200 ms in random order. In the icons the illuminated sector is shown in yellow; the black drawing represents the cell. Central sectors (**a**) always evoked depolarizations, but $[Ca^{2+}]$ increases only when the recording site was illuminated; note recording sites (marked by red and blue crosses) on roughly opposite sides of the soma. Integrated $[Ca^{2+}]$ responses (**b**) as a function of sector orientation for both recording sites (red and blue triangles) together with the electrical response (black circles). Error bars indicate s.d. ($n = 4$).

moved off the cell into the near surround. Switching off the bar in the far surround (about 1.2 mm from the cell) evoked a brief depolarizing response, presumably reflecting the release of the strong and long-ranging inhibition that has been reported for these cells¹⁸. The $[Ca^{2+}]$ response (bottom traces in each pair) began for all directions of motion as a slow, small depression later giving way to a recovery or even an increase above the resting $[Ca^{2+}]$ that was largest when the bar moved in a particular direction over the optical recording site.

Closer analysis of the responses to moving stimuli shows how strong the calcium direction selectivity is. To quantify directional selectivity we temporally integrated voltage and $[Ca^{2+}]$ changes (see

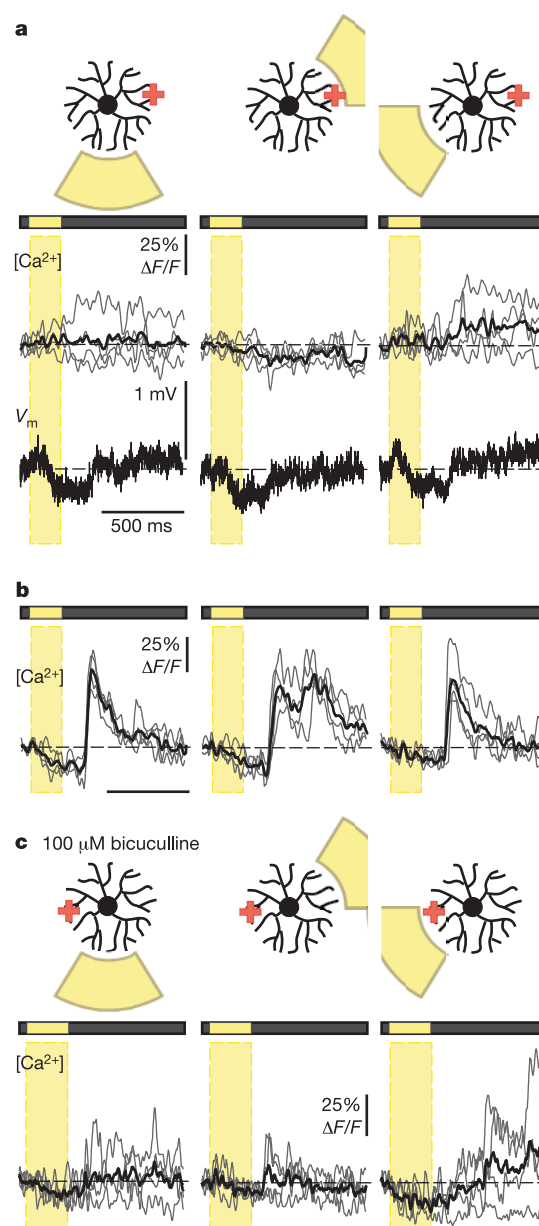


Figure 3 Surround stimulation evokes a decrease in $[Ca^{2+}]$ and occasionally transient increases after the end of the stimulus. Somatic voltage (V_m) and dendritic $[Ca^{2+}]$ in response to 'pie-wedge'-shaped 60°-sector stimuli flashed for 200 ms (400 ms in **c**) over the receptive field surround. $[Ca^{2+}]$ depression and hyperpolarization were almost always observed (**a** and **b**) but delayed $[Ca^{2+}]$ increase after stimulus-offset occurred only occasionally (**b**). Neither the $[Ca^{2+}]$ depression nor the delayed increase were abolished by the GABA_A blocker bicuculline (**c**).

Methods) and plotted the results as a function of motion direction using polar coordinates (Fig. 4c, d). The normalized vector sum of the responses points in the preferred direction with its length providing a measure (the directional preference index, DPI; see Methods) of the strength of the motion-direction discrimination. Preferred direction and DPIs (for the responses shown in Fig. 4a–d) are represented by arrows (resultant vectors) originating at the approximate recording sites on a schematic starburst cell (Fig. 4e).

The DPIs for voltage (V_m) and $[Ca^{2+}]$ responses to a drifting grating were 0.09 and 0.67, respectively. The corresponding DPIs for the moving bar were 0.08 and 0.40. The histogram of DPIs for all cells tested is shown in Fig. 4f. The DPI values (mean $\pm$ s.e.m) for all cells were 0.18 ± 0.02 ($n = 9$) and 0.43 ± 0.04 ($n = 12$) for V_m and dendritic $[Ca^{2+}]$ responses, respectively.

We next analysed how the locally preferred direction relates to cellular morphology (Fig. 5). In most cases $[Ca^{2+}]$ signals were

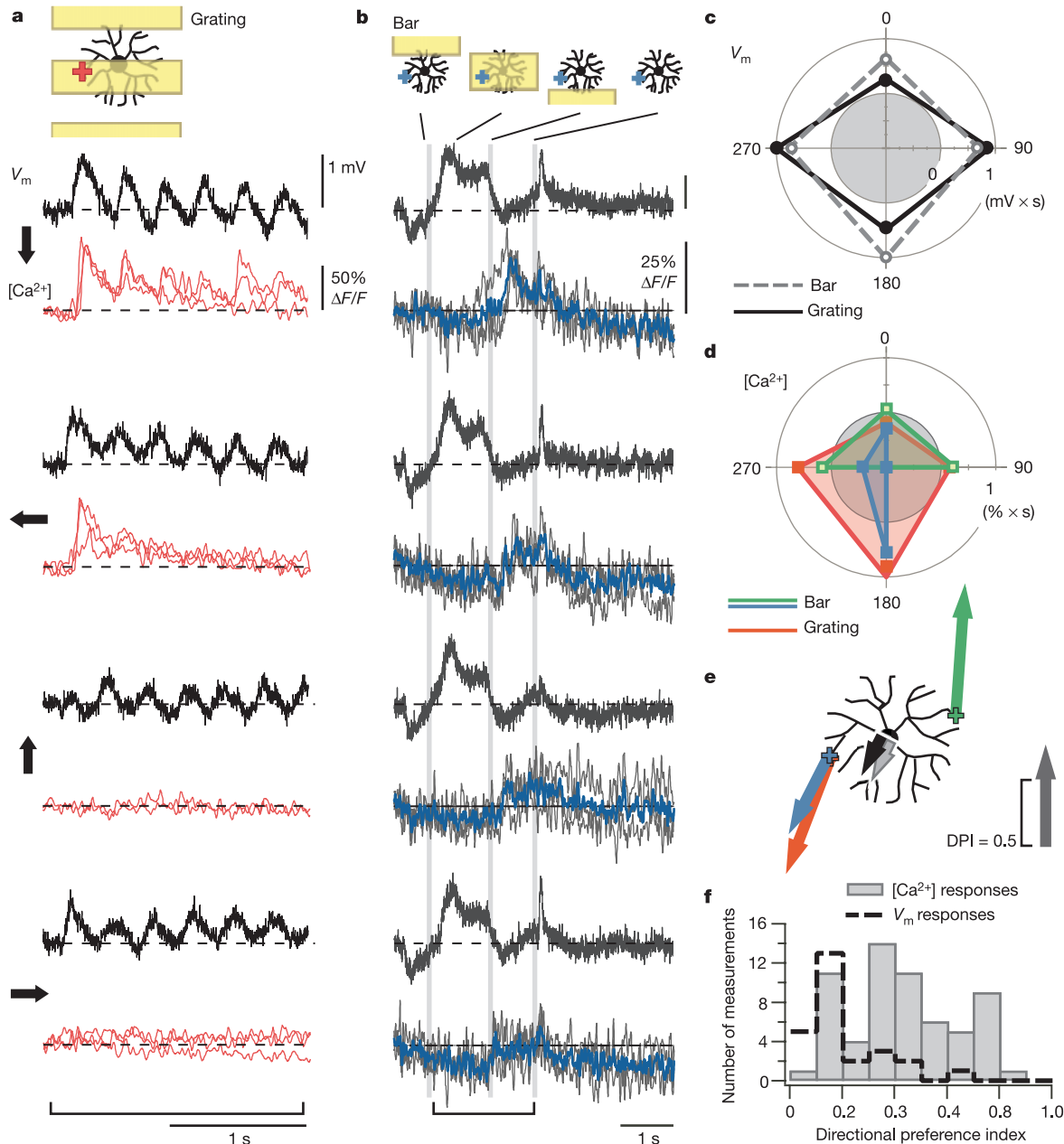


Figure 4 Moving stimuli evoke direction-selective calcium responses in dendrites. Somatic voltage (V_m) and dendritic $[Ca^{2+}]$ responses of a starburst cell to moving gratings (a) and moving bars (b); averages of three trials. Direction of movement indicated by arrows next to each pair of traces. Stimulus speed was 1.3 mm s^{-1} for the grating (period, $400 \mu\text{m}$; 82% contrast; mean intensity, about $700 \text{ photons s}^{-1} \mu\text{m}^{-2}$) and 1.0 mm s^{-1} for the bar ($500 \mu\text{m}$ wide; 87% contrast). The grating started moving 50 ms after the beginning of the traces (in a) and continued with constant velocity throughout the trace. Modulation of V_m (upper trace in each pair) did not depend strongly on motion direction either for the drifting grating (a) or for the moving bar (b). $[Ca^{2+}]$ changes (lower trace in each pair) in distal dendritic segments occurred only when the grating moved towards

180° or 270° (a). The bar moving towards 180° caused the largest $[Ca^{2+}]$ peak (b). Voltage (c) and $[Ca^{2+}]$ (d) responses integrated over the stimulus duration (marked by bracket below traces in a and b) and plotted as function of the direction of motion. Green trace in polar plot represents bar responses for second recording site located on the opposite side of the cell (traces not shown). e, Arrows on schematic cell represent the directionality vector (v). DPI values for $[Ca^{2+}]$ signals were 0.67 (red), 0.40 (blue), and 0.72 (green); for voltage responses 0.09 (black), and 0.08 (grey). Arrow origins indicate positions of recording sites. f, Histogram with DPI values for nine electrically recorded (dashed line; 26 measurements) and 12 optically recorded cells (62 measurements).

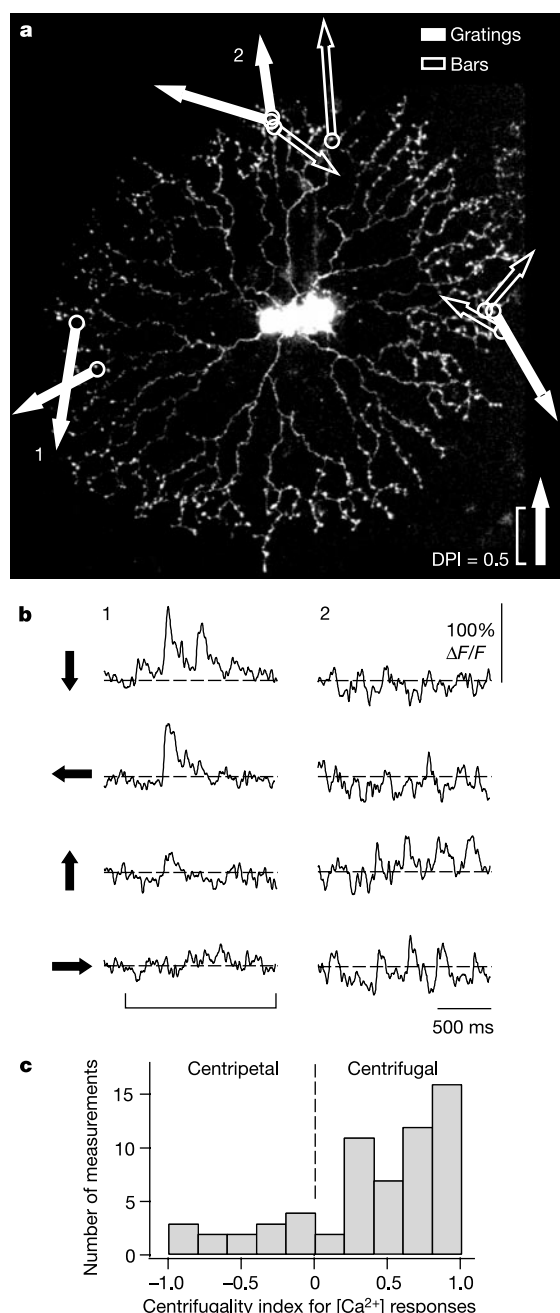


Figure 5 Preferred direction and cellular morphology. **a**, Preferred direction and DPI probed by bars (open arrows) or gratings (solid arrows) for various dendritic locations (circles at arrow origins) in one starburst cell. **b**, For two recording sites (arrows marked 1 and 2) the time traces of the grating $[Ca^{2+}]$ responses are shown. **c**, Distribution of centrifugality index (CI) values for all 12 cells tested (62 measurements).

larger for stimuli that moved outward from the soma, that is, for centrifugal directions of motion. This is apparent at several different recording sites for the cell in Fig. 5a. The directionality vectors at different dendritic sites did not, in general, point straight outward from the soma. We quantified the deviations from a strictly centrifugal line with a centrifugality index (CI) that ranged from +1 for perfect centrifugal preference to -1 for perfect centripetal preference (see Methods). For the cell in Fig. 5 the CI was 0.43 ± 0.18 ($n = 12$; DPI = 0.47 ± 0.07), indicating a clear preference for centrifugal motion; for all 12 cells tested the mean CI was 0.42 ± 0.06 (see histogram in Fig. 5c).

Blocking GABA_A receptors abolishes direction selectivity in DS

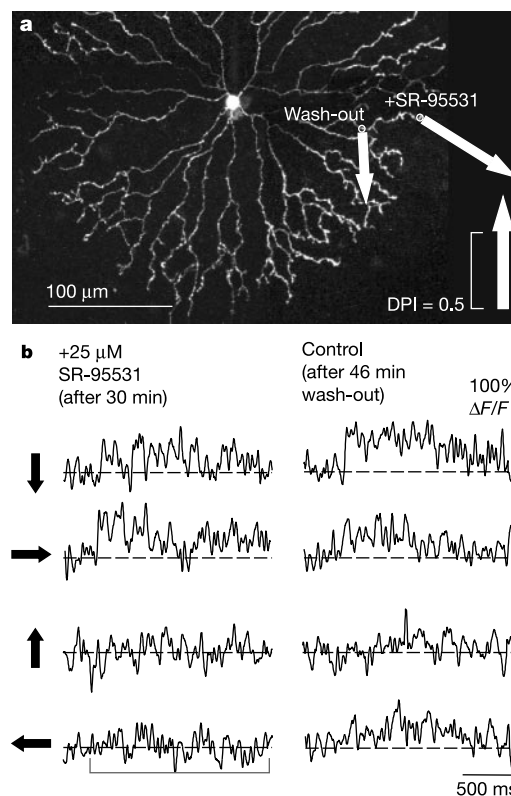


Figure 6 The GABA_A blocker SR-95531 caused increased spontaneous activity but no reduction of $[Ca^{2+}]$ signal direction selectivity. **a**, **b**, DPIs of 0.61 and 0.32 under GABA_A block (**a**) and after wash-out (**b**).

ganglion cells^{25,26} but not, as we found (Fig. 2c), the depression of $[Ca^{2+}]$ by surround illumination, so we investigated the effect of GABA_A receptor antagonists on the $[Ca^{2+}]$ responses to moving stimuli. We found that in the presence of either 100 μM bicuculline or 20–50 μM SR-95531 spontaneous activity increased, but evoked $[Ca^{2+}]$ signals remained direction-selective (DPI = 0.53 ± 0.04 ; CI = 0.45 ± 0.14 ; $n = 4$; Fig. 6a, b).

Somatic responses to circular wave stimuli

The $[Ca^{2+}]$ signals we have recorded show that dendritic sectors act as independent motion detectors with a direction of preference centrifugal from the soma. To test with somatic voltage recordings whether direction selectivity is also present in the electrical signal from each dendrite, we have to avoid cancellation of voltage contributions from different branches, which, for uniform motion, ‘see’ motion along preferred, neutral and null directions (depending on their orientation). Instead of uniform stimuli we therefore used expanding or contracting concentric waves (see Methods), which should excite each branch in the same sense. When comparing expanding and contracting stimuli we found very similar electrical responses (Fig. 7a), consisting of a net (direct current, d.c.) depolarization (expanding: 0.31 ± 0.07 mV; contracting: 0.42 ± 0.06 mV; ratios: 0.73 ± 0.16 , range: 0.14 to 1.19; $n = 6$; four cells) with a superimposed modulation (expanding: $V_{rms} = 0.33 \pm 0.07$ mV; contracting: $V_{rms} = 0.25 \pm 0.03$ mV; ratios: 1.28 ± 0.18 , range: 0.89 to 2.02). The $[Ca^{2+}]$ signals, on the other hand, were much larger for the expanding ($21 \pm 7\%$ $\Delta F/F$; $n = 7$; six cells) than for the contracting stimulus ($4 \pm 2\%$ $\Delta F/F$). Although there was no clear difference in either the net (integrated) electrical depolarization or the modulation amplitude we wondered if underlying differences might have been masked by strong proximal bipolar cell input⁵ driven by the modulation of the central

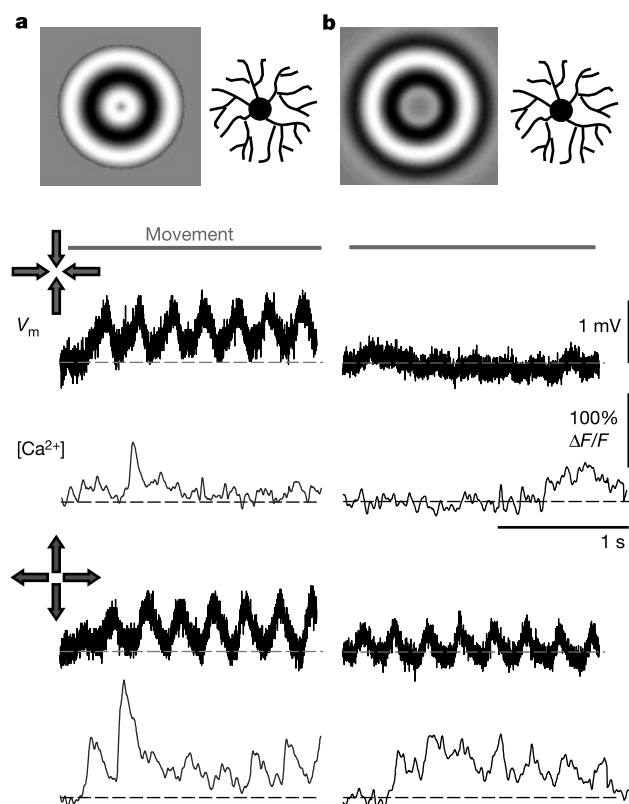


Figure 7 Responses to expanding and contracting waves. **a**, Somatic voltage (V_m) and dendritic $[Ca^{2+}]$ responses to stimuli comprising an expanding or contracting concentric sinusoidal wave (see Methods). **b**, A different cell excited by a stimulus similar to the one in **a** but with constant intensity in the central part of the stimulus and smoothed transitions between annulus and background. Icons next to stimuli represent relative size of cells.

intensity. To avoid exciting such input we modified the stimuli, with their centre intensity now kept constant (Fig. 7b). With these stimuli the net electrical depolarizations for expanding and contracting motion were small and not systematically different (expanding: 0.24 ± 0.06 ; contracting: 0.23 ± 0.05 mV; ratios: 0.97 ± 0.14 , range: 0.28 to 2.09; $n = 13$; six cells). The expanding stimulus evoked, however, a larger modulation of the membrane potential than the contracting stimulus (expanding: $V_{rms} = 0.25 \pm 0.06$ mV; contracting: $V_{rms} = 0.18 \pm 0.04$ mV; ratios: 1.52 ± 0.17 , range 1.08 to 3.38). Although the average ratio is significantly larger than one and we never found the contracting stimulus leading to a larger modulation, the difference between the expanding and contraction stimulus responses was sometimes small. The most likely reason is insufficient centring of the central grey area, leading to residual modulation of the proximal inputs.

The lack of direction-dependent differences in the d.c. (net) responses suggests that the net synaptic input current to the starburst cell is not direction-selective. The alternating current (a.c.) potential evoked by stimuli moving in the preferred direction might, however, generate a direction-selective net $[Ca^{2+}]$ signal by rectification, which occurs inevitably owing to the nonlinear dependence of calcium influx on the membrane voltage. This suggests that the dendritic branches of starburst cells perform the nonlinear computation required²⁷ for detection of directed motion.

Discussion

It has been proposed that direction selectivity in starburst cell dendrites could arise solely from intrinsic dendritic properties and the asymmetrical arrangement of its input and output synapses^{4,21}. In such a model, a stimulus moving in the centrifugal direction first

generates, when passing the proximal part of the dendrite, a depolarizing potential. As this potential spreads electrotonically to the distal dendrite it does so with a temporal delay that allows it to sum with the depolarization directly evoked there by the stimulus, which reaches that part of the cell at a slightly later time. A stimulus moving in the opposite direction excites the distal dendrite first and then later the proximal dendrite. This causes the direct response in the distal dendrite and the electrotonic potential that spreads there from its proximal site of generation to be out of phase and not effectively summed. Hence the peak depolarization in the distal dendrite, which is the region that contains the output synapses, would be smaller for centripetal than centrifugal movement. By the same reasoning, however, we expect strong summation at proximal sites of direct and electrotonically spreading signals for centripetal stimuli. This, however, is not what our circular wave stimuli show.

Thus a simple electrotonic model appears to be inconsistent with our observations. It therefore seems necessary to involve other elements such as active membrane conductances^{21,28}, cholinergic (acetylcholine-mediated) connections between amacrine cells^{5,29}, or inhibition⁴. Inhibitory inputs are present on starburst cells, as evident from physiological (Figs 2–4, and refs 24, 30) and histochemical³¹ data. Our observation that $[Ca^{2+}]$ signals remain direction-selective in the presence of GABA_A blockers shows that GABA_A receptors are not necessary. Although there is also a report that blocking GABA_C receptors in addition to GABA_A receptors does not further affect direction selectivity at the ganglion cell level²⁵, these results do not definitively rule out a role of GABA_C at the starburst cell level.

How is the directional motion information used after it has been generated in the starburst cell dendrite? Starburst amacrine cells contain and release GABA and acetylcholine (ACh)^{32–34}. There is no direct evidence that GABA release is direction-selective, but we do know that GABA is released upon membrane depolarization³⁵, probably using GABA transporters. Net GABA release might be direction-selective (via a rectification mechanism akin to that for calcium influx), because transporter flux is a strongly nonlinear function of membrane voltage³⁶ and dendritic membrane voltage is predominantly modulated by centrifugal motion (see Fig. 7b). Such a scenario is consistent with previous findings that DS ganglion cells receive asymmetrical inhibition⁸ that is mediated by GABA_A receptors^{25,26}. We also believe that the actual data in ref. 9 are consistent with asymmetrical inhibition originating presynaptically.

Direction-selective ACh release should occur since its release is vesicle-based and calcium-dependent³⁵. The strong supralinear (about 4th power³⁷) $[Ca^{2+}]$ dependence of vesicular release may thereby further enhance the substantial direction selectivity already present in the $[Ca^{2+}]$ responses. There have been, however, somewhat conflicting reports on the effect of ACh antagonists on direction selectivity in DS ganglion cells. Responses to moving-bar stimuli are uniformly reduced but direction discrimination appears to be undiminished^{7,38–40}. For responses to moving gratings and textures, direction selectivity appears to be abolished⁴¹ by ACh antagonists. These observations suggest that the cholinergic pathway is clearly not the only way for direction information to reach DS ganglion cells; yet cholinergic input (which only originates⁴² from starburst cells) to DS ganglion cells conveys directional information, at least for some stimulus types. Excitatory cholinergic inputs may cooperate with inhibitory inputs onto DS ganglion cells in a ‘push–pull’ configuration, with inhibitory input active during null-direction motion and excitatory inputs active during preferred-direction motion. Such a balanced arrangement would not only increase sensitivity but also serve to attenuate the ‘common mode’ response to mere changes in the mean intensity (see also ref. 43).

We still need to know why GABA blockers^{25,26} but not acetylcholine blockers^{7,38–40} abolish moving-bar direction selectivity. Moving bar stimuli impart a strong overall common-mode intensity change on the cell. Because removal of inhibition disrupts the

common-mode suppression afforded by the push–pull arrangement, it is conceivable that direction selectivity is masked by an increased spiking response to the overall intensity-change (which occurs in both preferred- and null-directions). Cholinergic block, on the other hand, removes excitatory drive. This would reduce the rate of spike production, but it would not eliminate directional selectivity by masking the difference in the responses to null and preferred stimuli.

Direction-selective release of ACh and/or GABA fits well within presynaptic models of direction selectivity⁴⁴. Rather specific wiring between starburst cell and DS ganglion cell dendrites is, however, required. If DS ganglion cells were to indiscriminately contact all starburst cell dendrites within their dendritic field they would receive direction-selective input for all directions, the sum of which would no longer be direction-selective (though it might still be motion sensitive⁷). Therefore, to use the direction-selective output of starburst cells DS ganglion cells have to contact selectively only that subset of dendrites that provides the ‘correct’ information. To achieve this, a DS ganglion cell dendrite has to receive excitatory cholinergic input from only those starburst cell branches with somata located on the preferred side of the ganglion cell (the side from which a preferred-direction stimulus approaches) and/or GABA-mediated inhibitory input from those branches with somata on the null side of the DS ganglion cell.

An asymmetric connectivity pattern might also account for the non-discriminatory zone found universally in DS ganglion cells⁴⁵ because, owing to the distal location of the outputs, motion information in the starburst cell is shifted peripherally. Hence, for inputs located at the preferred-direction edge of the DS ganglion-cell dendritic field the inhibitory outputs of the starburst cell would be shifted beyond the reach of the DS ganglion cell, whereas the excitatory outputs could connect, locally disrupting the balanced push–pull arrangement. Future measurement of excitatory and inhibitory synaptic currents in the DS ganglion cell for stimuli in the non-discriminatory zone should resolve this issue.

Whatever detailed mechanisms of direction detection in the retina eventually emerge, our results demonstrate that branch-specific dendritic computation does occur in the retina and that direction selectivity is computed at a stage presynaptic to ganglion cells with the starburst cell dendrites showing behaviour expected for a Reichardt correlation detector²⁷. □

Methods

Tissue and electrical recordings

Dark-adapted adult albino rabbits were deeply anesthetized with ketamine (Parke-Davis) and xylazine (Bayer), intramuscularly, and then killed with an intravenously applied overdose of pentobarbital (Narcoren; Merial). The eyes were enucleated and the retinas were dissected out under dim red illumination. A piece of mid-peripheral, flat-mounted retina was placed under the microscope and perfused with warmed (~36°C), oxygenized (95% O₂, 5% CO₂) Ames medium (Sigma). The tissue was counterstained with sulphorhodamine 101 dye (Sigma) that allowed us visually to identify with a high success rate (~90%) ON starburst amacrine cells, located in the ganglion cell layer, which were then penetrated using sharp micro-electrodes (filled with 0.2 M K-acetate; 100–250 MΩ; borosilicate with filament, of outer diameter, 1.0 mm and inner diameter 0.58 mm; Clark, Harvard Apparatus, pulled on a P2000 laser puller, Sutter Instruments). First, the voltage (V_m) response to the visual stimuli was recorded, and the cell was filled with the calcium indicator Oregon Green 488 BAPTA-1 (pipette concentration 5 mM in 0.2 M K-acetate; Molecular Probes). Then, the electrode was carefully retracted and the cell was given about 20–30 min to recover in sulphorhodamine-free medium before calcium imaging started. The GABA_A receptor antagonists bicuculline and SR-95531 were purchased from Sigma and applied with the extracellular medium.

Multi-photon calcium imaging

An upright multi-photon microscope (custom in-house design, software developed at Bell Labs by R. Stepnoski and modified at the Max Planck Institute for Medical Research by M. Müller) was used to image calcium signals from selected dendrites¹⁶. A mode-locked Ti:sapphire laser (Mira-900; Coherent) tuned to 930 nm was used as an excitation source. The scanning laser beam caused only moderate and quickly adapting photoreceptor excitation, allowing us to visually stimulate and optically record calcium signals simultaneously. Filters in the stimulation light path (600 BP 20) and in front of the detector (520 BP 30) ensured that stimulus light did not interfere with fluorescence detection. Image series (64 × 8 pixel images at 62.5 Hz) or line scans (64 pixel lines at

500 Hz) were acquired from short dendritic segments (using water immersion lenses from Zeiss, 63 × 0.9 NA (numerical aperture); or Leica, 60 × 0.9 NA) and analysed off-line (IgorPro; Wavemetrics). While the observed changes in fluorescence suggest changes in [Ca²⁺] of several hundred nM no attempt was made at absolute quantification because no information is available on the endogenous calcium buffer capacity of starburst dendrites.

Light stimulation

The light stimuli, generated by custom-written software, were output through a video projector (Astrobeam 530; A + K, Germany), and focused through a substage condenser lens (Zeiss 0.32 NA) onto the photoreceptors. The stimulus intensity ranged from about 130 to 2,000 photons s⁻¹ μm⁻², roughly covering the mesopic range; stimulus contrasts were between 70 and 87%. The velocity for moving stimuli (bars, gratings) ranged from 500 to 1,300 μm s⁻¹. ‘Pie-wedge’-shaped sector stimuli (see Figs 2, 3) illuminated 60° sectors of either the cell’s dendritic arbor (central sectors) or the near surround (annular surround sectors) for 200 or 400 ms. ‘Bull’s eye’ stimuli (see Fig. 7) consisted of expanding or contracting concentric sinusoidal waves that were centred on the cell and scaled to cover the dendritic field (about 1.5 cycles per cell’s radius moving at 3–4 cycles s⁻¹). Stimuli were repeated 3–6 times in random order.

Data analysis

Fluorescence (F) data were spatially averaged over selected regions of interest (ROI) and temporally filtered (box car; image series: 11 frames = 176 ms; line scans, 25 ms).

After background subtraction $\Delta F/F$ was calculated for each stimulus presentation. To quantify V_m and $\Delta F/F$ responses the area under the data traces for a defined interval (between t_1 and t_2) was calculated:

$$a = \int_{t_1}^{t_2} (\Delta F/F) dt$$

Several results for the same stimulus condition were averaged. A directionality vector was calculated as follows:

$$\mathbf{v} = [(a_{+x} - a_{-x})\mathbf{e}_x + (a_{+y} - a_{-y})\mathbf{e}_y] / (a_{+x} + a_{-x} + a_{+y} + a_{-y})$$

where $\mathbf{e}_{x,y}$ are unit vectors. For each [Ca²⁺] measurement location and for the membrane voltage the value of the directional preference index

$$\text{DPI} = |\mathbf{v}| = \sqrt{(a_{+x} - a_{-x})^2 + (a_{+y} - a_{-y})^2} / (a_{+x} + a_{-x} + a_{+y} + a_{-y})$$

was then calculated. The DPI, for which $0 \leq \text{DPI} \leq 1$ holds, is a measure of directional asymmetry but does not indicate the direction that is preferred. We therefore also used a centrifugality index ($\text{CI} = \cos(\alpha)$), with the angle α between the radial direction and $\mathbf{v}$), which indicates whether centrifugal ($\text{CI} > 0$) or centripetal ($\text{CI} < 0$) motion was preferred. Voltage responses to ‘Bull’s eye’ stimuli were fitted with a sum of a constant plus a sinusoidal wave to quantify d.c. depolarization and modulation. For ‘Bull’s eye’-evoked [Ca²⁺] responses the average $\Delta F/F$ was determined.

Received 2 April; accepted 17 June 2002; doi:10.1038/nature00931.

Published online 4 August 2002.

- Barlow, H. B., Hill, R. M. & Levick, W. R. Rabbit retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol.* **173**, 377–407 (1964).
- Vaney, D. I., He, S., Taylor, W. R. & Levick, W. R. in *Motion Vision—Computational, Neural, and Ecological Constraints* (eds Zanker, J. M. & Zeil, J.) (Springer, Berlin, 2001).
- Masland, R. H., Mills, J. W. & Cassidy, C. The functions of acetylcholine in the rabbit retina. *Proc. R. Soc. Lond. B* **223**, 121–139 (1984).
- Borg-Graham, L. J. & Grzywacz, N. M. in *Single Neuron Computation* (eds McKenna, T., Zornetzer, S. F. & Davis, J. L.) 347–375 (Academic, London, 1992).
- Famiglietti, E. V. Synaptic organization of starburst amacrine cells in rabbit retina: analysis of serial thin sections by electron microscopy and graphic reconstruction. *J. Comp. Neurol.* **309**, 40–70 (1991).
- Tsuchi, M. & Masland, R. H. The shape and arrangement of the cholinergic neurons in the rabbit retina. *Proc. R. Soc. Lond. B* **223**, 101–191 (1984).
- He, S. & Masland, R. H. Retinal direction selectivity after targeted laser ablation of starburst amacrine cells. *Nature* **389**, 378–382 (1997).
- Borg-Graham, L. J. The computation of directional selectivity in the retina occurs presynaptic to the ganglion cell. *Nature Neurosci.* **4**, 176–183 (2001).
- Taylor, W. R., He, S., Levick, W. R. & Vaney, D. I. Dendritic computation of direction selectivity by retinal ganglion cells. *Science* **289**, 2347–2350 (2000).
- Yoshida, K. *et al.* A key role of starburst amacrine cells in originating retinal directional selectivity and optokinetic eye movement. *Neuron* **30**, 771–780 (2001).
- Euler, T. & Denk, W. Dendritic processing. *Curr. Opin. Neurobiol.* **11**, 415–422 (2001).
- Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
- Ehinger, B., Zucker, C. L., Bruun, A. & Adolph, A. *In vivo* staining of oligodendroglia in the rabbit retina. *Glia* **10**, 40–48 (1994).
- Miller, R. F. *et al.* Structure and functional connections of presynaptic terminals in the vertebrate retina revealed by activity-dependent dyes and confocal microscopy. *J. Comp. Neurol.* **437**, 129–155 (2001).
- Brandon, C. Cholinergic neurons in the rabbit retina: dendritic branching and ultrastructural connectivity. *Brain Res.* **426**, 119–130 (1987).
- Denk, W. & Detweiler, P. B. Optical recording of light-evoked calcium signals in the functionally intact retina. *Proc. Natl Acad. Sci. USA* **96**, 7035–7040 (1999).
- Bloomfield, S. A. Relationship between receptive and dendritic field size of amacrine cells in the rabbit retina. *J. Neurophys.* **68**, 711–725 (1992).
- Taylor, W. R. & Wässle, H. Receptive field properties of starburst cholinergic amacrine cells in the rabbit retina. *Eur. J. Neurosci.* **7**, 2300–2321 (1995).

19. Peters, B. N. & Masland, R. H. Responses to light of starburst amacrine cells. *J. Neurophysiol.* **75**, 469–480 (1996).
20. Miller, R. F. & Bloomfield, S. A. Electroanatomy of a unique amacrine cell in the rabbit retina. *Proc. Natl Acad. Sci. USA* **80**, 3069–3073 (1983).
21. Poznanski, R. R. Modelling the electrotonic structure of starburst amacrine cells in the rabbit retina: a functional interpretation of dendritic morphology. *Bull. Math. Biol.* **54**, 905–928 (1992).
22. Taylor, W. R. & Wässle, H. Receptive field properties of starburst cholinergic amacrine cells in the rabbit retina. *Eur. J. Neurosci.* **7**, 2308–2321 (1995).
23. Pan, Z. H. & Lipton, S. A. Multiple GABA receptor subtypes mediate inhibition of calcium influx at rat retinal bipolar cell terminals. *J. Neurosci.* **15**, 2668–2679 (1995).
24. Linn, D. M. & Massey, S. C. GABA inhibits ACh release from the rabbit retina: a direct effect or feedback to bipolar cells? *Vis. Neurosci.* **8**, 97–106 (1992).
25. Massey, S. C., Linn, D. M., Kittila, C. A. & Mirza, W. Contributions of GABAA receptors and GABAC receptors to acetylcholine release and directional selectivity in the rabbit retina. *Vis. Neurosci.* **14**, 939–948 (1997).
26. Wyatt, H. J. & Daw, N. W. Specific effects of neurotransmitter antagonists on ganglion cells in rabbit retina. *Science* **191**, 204–205 (1976).
27. Reichardt, W. in *Sensory Communication* (ed. Rosenblith, W. A.) 377–390 (MIT Press/Wiley, New York, 1961).
28. Velte, T. J. & Miller, R. F. Spiking and nonspiking models of starburst amacrine cells in the rabbit retina. *Vis. Neurosci.* **14**, 1073–1088 (1997).
29. Keyser, K. T. *et al.* Amacrine, ganglion, and displaced amacrine cells in the rabbit retina express nicotinic acetylcholine receptors. *Vis. Neurosci.* **17**, 743–752 (2000).
30. Zhou, Z. J. & Fain, G. L. Neurotransmitter receptors of starburst amacrine cells in rabbit retinal slices. *J. Neurosci.* **15**, 5334–5345 (1995).
31. Wässle, H., Koulen, P., Brandstätter, J. H., Fletcher, E. L. & Becker, C. M. Glycine and GABA receptors in the mammalian retina. *Vision Res.* **38**, 1411–1430 (1998).
32. Masland, R. H., Mills, J. W. & Hayden, S. A. Acetylcholine-synthesizing amacrine cells: identification and selective staining by using radioautography and fluorescent markers. *Proc. R. Soc. Lond. B* **223**, 79–100 (1984).
33. Brecha, N., Johnson, D., Peichl, L. & Wässle, H. Cholinergic amacrine cells of the rabbit retina contain glutamate decarboxylase and gamma-aminobutyrate immunoreactivity. *Proc. Natl Acad. Sci. USA* **85**, 6187–6191 (1988).
34. Vanev, D. I. & Young, H. M. GABA-like immunoreactivity in cholinergic amacrine cells of the rabbit retina. *Brain Res.* **438**, 369–373 (1988).
35. O'Malley, D. M., Sandell, J. H. & Masland, R. H. Co-release of acetylcholine and GABA by the starburst amacrine cells. *J. Neurosci.* **12**, 1394–1408 (1992).
36. Schwartz, E. A. Depolarization without calcium can release gamma-aminobutyric acid from a retinal neuron. *Science* **238**, 350–355 (1987).
37. Dodge, F. A. Jr & Rahamimoff, R. Co-operative action a calcium ions in transmitter release at the neuromuscular junction. *J. Physiol.* **193**, 419–432 (1967).
38. Ariel, M. & Daw, N. W. Pharmacological analysis of directionally sensitive rabbit retinal ganglion cells. *J. Physiol.* **324**, 161–185 (1982).
39. Grzywacz, N. M., Tootle, J. S. & Amthor, F. R. Is the input to a GABAergic or cholinergic synapse the sole asymmetry in rabbit retinal directional selectivity? *Vis. Neurosci.* **14**, 39–54 (1997).
40. Kittila, C. A. & Massey, S. C. Pharmacology of directionally selective ganglion cells in the rabbit retina. *J. Neurophysiol.* **77**, 675–689 (1997).
41. Grzywacz, N. M., Merwine, D. K. & Amthor, F. R. Complementary roles of two excitatory pathways in retinal directional selectivity. *Vis. Neurosci.* **15**, 1119–1127 (1998).
42. Masland, R. H. & Mills, J. W. Autoradiographic identification of acetylcholine in the rabbit retina. *J. Cell Biol.* **83**, 159–178 (1979).
43. Borst, A. & Egelhaaf, M. Direction selectivity of blowfly motion-sensitive neurons is computed in a two-stage process. *Proc. Natl Acad. Sci. USA* **87**, 9363–9367 (1990).
44. Vanev, D. I., Collins, S. P. & Young, H. M. in *Neurobiology of the Inner Retina* (eds Weiler, R. & Osborn, N. N.) 157–168 (Springer, Berlin, 1989).
45. Barlow, H. B. & Levick, W. R. The mechanism of directionally selective units in rabbit's retina. *J. Physiol.* **178**, 477–504 (1965).

Acknowledgements

We thank A. Borst and D. Margolis for useful discussions, R. Friedrich for comments on the manuscript, and M. Müller, J. Tritthart and J. Sawinski for technical help.

Competing interests statement

The authors declare competing financial interests: details accompany the paper on *Nature's* website (<http://www.nature.com/nature>).

Correspondence and requests for materials should be addressed to T.E. (e-mail: teuler@mpimf-heidelberg.mpg.de).

A wind origin for Titan's haze structure

P. Rannou*, F. Hourdin† & C. P. McKay‡

* Service d'Aéronomie, B102, Université de Paris 6/Université de Versailles-St-Quentin, 4 place Jussieu, 75252 Paris Cedex, France

† Laboratoire de Météorologie Dynamique, Université de Paris 6, 4 place Jussieu, 75252 Paris Cedex, France

‡ NASA-Ames Research Center, Moffett Field, California 94035, USA

Titan, the largest moon of Saturn, is the only satellite in the Solar System with a dense atmosphere. Titan's atmosphere is mainly nitrogen with a surface pressure of 1.5 atmospheres and a temperature of 95 K (ref. 1). A seasonally varying² haze, which appears to be the main source of heating and cooling that drives atmospheric circulation^{3,4}, shrouds the moon. The haze has numerous features that have remained unexplained. There are several layers⁵, including a 'polar hood'^{6–8}, and a pronounced hemispheric asymmetry². The upper atmosphere rotates much faster than the surface of the moon^{9,10}, and there is a significant latitudinal temperature asymmetry at the equinoxes^{11,12}. Here we describe a numerical simulation of Titan's atmosphere, which appears to explain the observed features of the haze. The critical new factor in our model is the coupling of haze formation with atmospheric dynamics, which includes a component of strong positive feedback between the haze and the winds.

Earth-based observations have long indicated the presence of CH₄ in Titan's atmosphere and the possibility of a rich organic chemistry. The complexity of Titan's atmospheric photochemistry was confirmed by Voyager observations in 1980 which revealed the presence of nine gaseous organic molecules other than CH₄ (refs 13, 14). In addition, laboratory simulations using gas mixtures identical to Titan's atmosphere have shown that the solid organic material produced in the laboratory has the same optical properties as the haze determined from Titan's geometric albedo^{15,16}. Radiative transfer calculations show that this haze is the dominant absorber of sunlight in Titan's atmosphere, particularly in the ultraviolet and blue regions of the spectrum, with the result that only 10% of the solar flux at the top of Titan's atmosphere reaches the surface^{16–18}. A significant fraction of sunlight (40%) is absorbed by the haze in the stratosphere, creating an antigreenhouse effect¹⁸. In addition, the haze is also the dominant source of infrared cooling in the stratosphere¹⁶, and therefore it is not surprising that the stratospheric circulation is dominated by radiative balance due to the haze.

Timescale analysis has shown that the observed seasonal variation is not due to changes in haze production rate¹⁹ as originally suggested^{20,21}. Titan's year lasts about 30 terrestrial years, and microphysical models show that the timescales for aerosols to grow and fall to the troposphere is several decades of Titan years^{19,22}. Thus, haze seasonal variations must be due to dynamical effects^{4,23,24}. Similarly, one-dimensional microphysical models suggested that the detached haze was due to dynamical transport of small particles to higher elevation²³.

Pioneer simulations of Titan's atmospheric dynamics with a general circulation model³ succeeded in producing a strong super-rotation with prograde equatorial zonal winds of about 100 m s^{−1} near the 1-mbar level. However, these simulations were strongly underestimating latitudinal thermal contrasts in the same altitude range. It was suggested at that time that the model deficiency could be due to the lack of latitudinal and seasonal changes in the distribution of chemical species and haze. It was also thought that a simulation that correctly couples the haze transport and the diabatic heating that drives the stratospheric circulation would be necessary to self-consistently explain the observed winds, the thermal contrasts and the main features of the haze.

To investigate the nature of the interactions between Titan's haze and dynamics we developed a coupled fractal microphysics and two-dimensional circulation model, as described in the Methods section. In this version of the model, we need to parameterize the barotropic instabilities which contribute to dissipation of the wind latitudinal gradients. This is a purely three-dimensional effect of dynamics, which we account for by a simple diffusion operator (see Methods section). Aside from this, the equations of motion are explicitly solved by a numerical approach. Table 1 presents an overview of the model prediction compared to observations. The results for the haze distribution at the epoch just before the Voyager observations are shown in Fig. 1.

The meridional circulation in the model consists of a summer-pole-to-winter-pole cell which persists for about 80% of the time between consecutive equinoxes. The reversal of the circulation near equinoxes produces a transitory equator-to-poles cell system³. The aerosols are produced around 450 km (refs 25, 26) within a zone which has a vertical extent of about one scale height ($H = 40$ km). There, the meridional winds are of the order of $v = 2.5$ m s^{−1}, whereas the aerosol settling velocities do not exceed $w = 10^{-2}$ m s^{−1}. Consequently, the time to move over an horizontal length equal to Titan's radius ($R_t/v \approx 10^6$ s) is three to four times shorter than the time to fall one atmospheric scale height ($H/w \approx 3 \times 10^6$ s). These times may also be compared to the characteristic growth time for aerosols to reach the monomer radius for the fractal particles ($\tau \approx 10^6$ s) (ref. 22). As a result, the particles

Table 1 Model predictions and observations

	Observation	Model	Comments
North-south asymmetry			
Green	17–20%, $k = 0.942$	20%, $k \approx 0.90$ –0.95	Depends on the haze vertical profile, and thus, on the aerosol structure and on a possible vertical diffusion
Blue	17–20%, $k = 0.887$	12%, $k \approx 0.75$	
Violet	17–20%, $k = 0.778$	0%, $k = 0.60$	
Detached haze	Maximum at 370 km Gap at 330 km Extinction $a = 1.7 \times 10^{-7}$ m ^{−1}	Maximum at 420 km Gap at 350 km $a = 5 \times 10^{-7}$ m ^{−1}	Detached haze in the right place, but too optically thick owing to poor knowledge about actual aerosol material and structure
Northern winter polar hood	Northward 65° N	Northward 65° N	
Latitudinal temperature			
At 1 mbar	From 169 K near equator to 152 K at 60° N and to 167 K at 55° S ($\Delta T \approx 17$ K)	Match* the relative contrasts for $L_s \uparrow = 0^\circ$ within 4 K, except near 50° N at 0.4 mbar where $\Delta T \approx 7$ K	Discrepancies in relative ΔT and time shift may be due to the horizontal diffusion scheme
At 0.4 mbar	From 179 K near equator to 163 K at 60° N and to 177 K at 55° S ($\Delta T \approx 17$ K)	No marked north-south asymmetry at $L_s = 9^\circ$	
Zonal winds at 1 mbar	High latitude jet (70° S) ≈ 150 m s ^{−1} and tropical winds ≈ 80 –100 m s ^{−1}	Latitude jet at 50° S, jet winds at ≈ 150 m s ^{−1} and tropical winds ≈ 120 m s ^{−1}	Shift in the location of the jet may also be due to the horizontal diffusion scheme

* Accounting for $\pm 5^\circ$ uncertainties on latitude of data.

† The solar longitude, L_s , indicates the location of a planet on its orbit around the Sun from the northern spring equinox $L_s = 0^\circ$. L_s is between 0° and 360° .

are driven to the winter pole by the pole-to-pole cell following a rather horizontal path. They grow by coagulation and settle near the pole. Thereafter, particles are carried by latitudinal winds over the rest of the planet when the strong polar downwelling branch is progressively replaced by an ascending branch just after equinoxes. At this time, the top of the polar haze (≈ 350 km) is globally lower than the detached haze (the top of which is located at 420 km and its bottom at 400 km) (Fig. 1). Polar aerosols are then redistributed about one scale height below the detached haze. At the same time, the detached haze is broken by the transitory equator-to-pole cell system, until the opposite pole-to-pole cell begins.

The transfer of haze to the winter pole which takes place most of the time has important consequences for Titan's atmosphere and explains several puzzling features of the haze. First, the poleward motion of the haze in the production zone and the redistribution of the haze a scale height lower gives a consistent mechanism for the formation of the detached haze layer by producing a zone where aerosol extinction is smaller than the surrounding altitudes. Aerosol size distributions is widely polydisperse ($dN/d\log r \propto r^{-a}$ with $a > 1$) inside the formation zone (detached haze). In this region, aerosols grow spherically from macromolecules to submicrometre particles. Outside the detached haze, the distribution is monomodal and aerosols grow as fractal aggregates^{25,26}. Thus, considering the motion of the aerosols as explained above, as well as Fig. 1, we find that aerosols leave the aerosol production zone near the pole and then start to grow fractally.

Secondly, the pole-to-pole circulation is also responsible for the observed hemispheric contrast. The haze material is produced at about 400 km and the aerosol mass mixing ratio decreases with altitude below this level. In the summer hemisphere, the ascending branch (between about the equator and the polar circle) carries up air from lower layers that is relatively poorly loaded with aerosols. In the winter hemisphere, the aerosols from the production zone are carried horizontally from the summer to the winter hemisphere, producing the detached haze layer, and sink near the winter pole. This tends to continuously impoverish the summer hemisphere in aerosols. Because atmospheric gas is bright (compared to the haze) in the visible, the summer hemisphere then appears brighter than

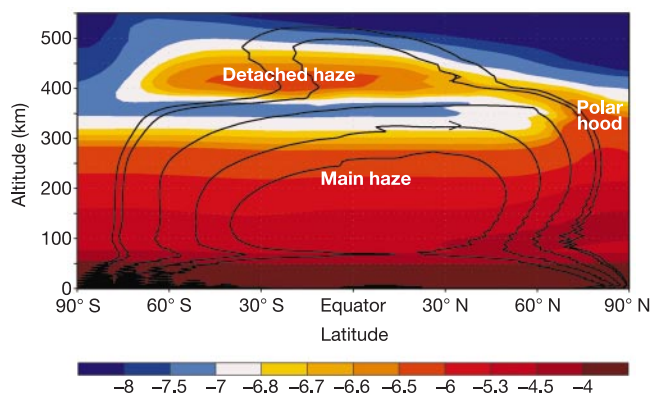


Figure 1 Two-dimensional planetary distribution of haze scaled extinction. The scaled extinction coefficient ($\beta \times \omega \times P(25^\circ)$) is shown as a function of the latitude and the pressure for the time $L_s = 353^\circ$ (less than a quarter of a season before actual Voyager time). Here β is the haze extinction, ω is the mean single scattering albedo and $P(25^\circ)$ is the mean phase function taken at a scattering angle 25° . The colour scale shows the $\log_{10}$ of the scaled extinction. This physical quantity directly shows the scattering efficiency of each level for a phase angle of $\Phi = 155^\circ$. The detached haze appears as a secondary layer at ≈ 400 km overlying the main haze (below 300 km) from one pole to the other. The haze is preferentially accumulated near the pole. The wind direction is illustrated by the stream function integrated from $L_s = 274^\circ$ to $L_s = 353^\circ$, where L_s is the solar longitude (northern spring equinox occurs at $L_s = 0^\circ$). The winds are from the south polar region (summer) toward the north polar region (winter).

the winter hemisphere, as observed by the Voyager spacecraft (Fig. 2). For wavelengths larger than about 650 nm, the haze becomes brighter than the atmospheric gases and surface below. Thus the summer hemisphere is darker than the winter hemisphere. The size of the aerosols also vary from one hemisphere to the other—by less than 20% in the main haze—but their properties only weakly depend on their size. Thus it does not affect the haze optical properties as much as concentration differences.

A third effect of this circulation pattern is a quasi-permanent accumulation of haze in the polar regions ($\Theta > 60^\circ$) (Fig. 1). The haze is typically three to five times optically thicker near the poles than in low latitudes. This accumulation is more marked at the winter pole, to which haze material is dynamically transported. In the summer pole the haze accumulation relaxes by sedimentation and removal by winds. The polar enhancement of the haze is recognizable throughout the year in the model, although observations do not clearly indicate such an increase around the summer pole^{5,6,8}. A lack of horizontal dissipation of the aerosol concentration gradients by transient eddies may account for this difference. The winter accumulation can be identified as the dark north polar hood seen by Voyager². Half a Titan year later (1995–2000), a southern polar hood is also visible in Earth-based infrared observations^{6,8}.

The permanent accumulation of haze in the polar regions produces a stronger infrared cooling during the polar night ($\Theta > 60^\circ$) than in the previous fixed-haze model³. In other regions where the haze is in sunlight, the infrared cooling rate and solar (visible) heating rate are roughly balanced, offsetting any variation in the haze. Thus, the net effect is that the accumulation of haze at the poles reinforces, on an annual basis, the equator-to-pole thermal contrasts, compared to the case for a homogeneous haze layer. This effect is maximized because the haze accumulation reaches its maximum at the winter pole.

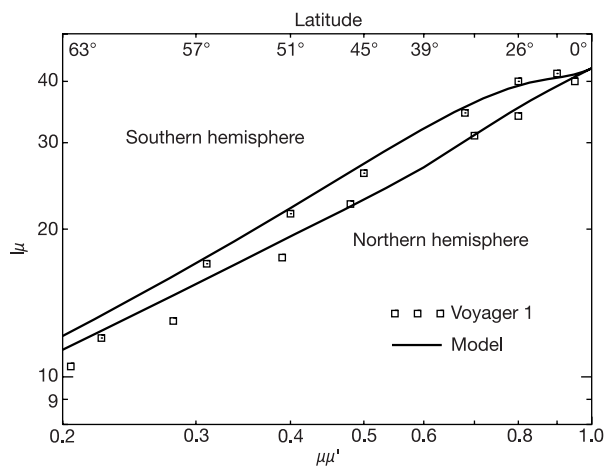


Figure 2 Comparison of the reflected intensity north-south profile between observation and model. The reflected intensity on Titan in green ($0.52\text{--}0.60\ \mu\text{m}$) observed along a meridian passing close to the subsolar point and the subspacecraft point compared with the model results at $0.55\ \mu\text{m}$ (ref. 2). Empty and dotted squares are data for the northern and southern hemispheres respectively. The intensity data (I) is plotted as product μ/μ' as a function of product $\mu\mu'$, where μ is the cosine of the observer zenith angle and μ' is the cosine of the solar zenith angle. The axis system at the top shows the south or north latitudes on the planet that correspond to the values of $\mu\mu'$ shown at the bottom. This axis system reveals any intensity behaviour which follows a Minnaert law $I(\mu) = [\mu\mu']^k I_0$. The exponent k characterizes the centre ($\mu = 1$) to limb ($\mu = 0$) darkening of the intensity. Observation shows that reflected intensity in the southern hemisphere is $\approx 25\%$ higher than in the northern hemisphere. The model is able to reproduce the observed north-south asymmetry at the Voyager epoch ($L_s = 9^\circ$) with the same centre to limb slope.

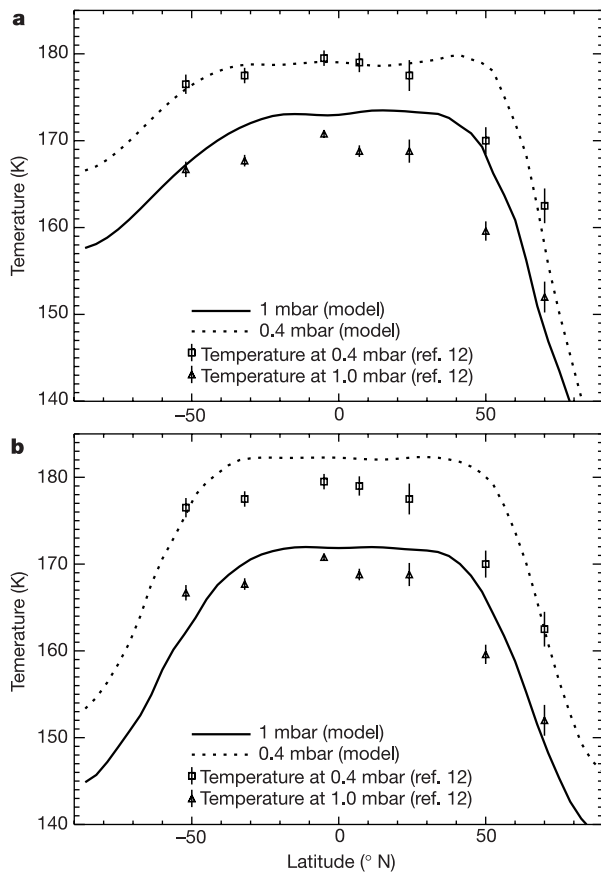


Figure 3 Comparison of latitudinal temperature profiles between observation and model. The latitudinal temperature profile at 1 mbar and 0.4 mbar retrieved from Voyager infrared data¹² and computed by the model at $L_s = 0^\circ$ (northern spring equinox, **a**) and $L_s = 9^\circ$ (Voyager epoch, **b**). Temperature profiles are close to observation, but in the model, the observed temperature asymmetry is better matched 1/8th of the season before the actual Voyager time. The model and data must be compared keeping in mind that observational temperature information is retrieved from an infrared radiative transfer model, and is thus also somewhat model dependent. In particular, temperature data have been retrieved^{11,12} from the same infrared measurements, and differences between the two set of results are of the same order, or even larger, than differences between our model and the retrieved observations presented here.

The response of atmospheric motions to the forcing produced by haze accumulation at the winter pole is that the strength of the mean meridional circulation is increased by a factor of two relative to a fixed haze model. This in turn reinforces accumulation of haze in polar regions, completing the basis of a strong positive feedback. The radiative forcing by the latitudinal contrast in the haze is of the same order as those produced by the latitudinal and seasonal varying insolation.

The stronger cooling in the polar regions also results in a larger equator-to-pole temperature contrast between 10 and 1 mbar and is associated through cyclostrophic balance with a stronger winter jet than in the fixed haze model³. This explains the latitudinal temperature gradient at 0.4 and 1 mbar observed by Voyager¹² (Fig. 3) and the strong polar jet⁹ (Fig. 4). □

Methods

We simulate the haze and atmospheric dynamics on Titan using a general circulation Model (GCM)³ that we have coupled to an eulerian microphysical model²⁶. The dynamical core of the GCM is based on a finite difference formulation of the primitive equations of meteorology. We use an axially symmetric version of the model which applies well to the Titan case. The probable weak influence of the diurnal cycle in the low stratosphere, the small longitudinal gradients compared to the latitudinal gradients and the strong zonal winds support this simplification. However, non-axially symmetric planetary waves produced by the barotropic instabilities are important in the creation and maintenance of

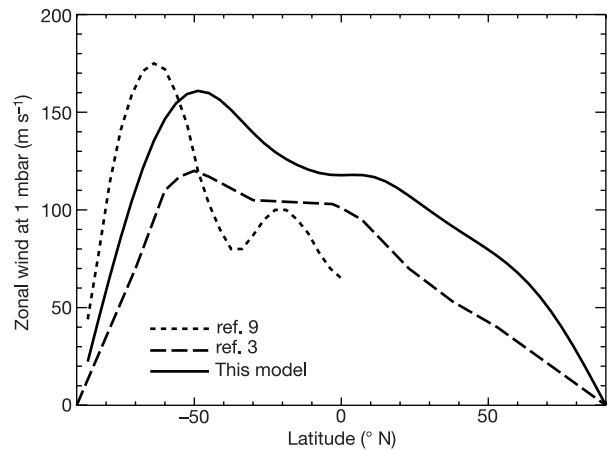


Figure 4 Comparison of latitudinal zonal wind profiles between observation and model. The latitudinal zonal wind profile at 1 mbar and $L_s = 106^\circ$ in our model (solid line) and for a GCM with a fixed haze³ (long dashed line) are compared to the zonal wind profile retrieved from the observation of occultation of 28 Sagittarii (short dashed line) at the same period⁹. The retrieved profile is deduced from a polynomial model of the atmospheric shape which best fits the occultation observation. The wind profile is derived from the two first polynomial terms only, and thus it essentially represents a rough indication of the magnitude and the shape of the actual latitudinal zonal wind profile. Our model is able to generate a strong zonal wind as observed, with a vigorous winter stratospheric jet which is attributed to the accumulation of haze near the winter poles.

the stratospheric super-rotation³. In the two-dimensional version of the GCM, the latitudinal transport by barotropic eddies must be parameterized. This approach was validated through a comparison with the results of the three-dimensional version of the model of ref. 3 and successfully applied by ref. 27 to study the transport of chemical species in Titan's atmosphere.

The eulerian microphysical model²⁶ was included in the GCM to account for the production, coagulation and sedimentation of fractal aggregate aerosols. To determine the optical properties of the fractal aggregates, we use a semi-empirical code²⁸ of scattering. The concentration and hence heating rate of the haze varies in position and time in response to transport and microphysical processes. This is the main difference from the previous models^{3,4}, which assumed a homogeneous or prescribed haze distribution of spherical aerosols.

The model runs with 48 points equally spaced in latitude ($\delta Y = 170$ km) and 55 levels of altitude roughly equally spaced in the log of pressure ($\delta Z \approx 10$ km in the stratosphere). The main free parameter for the dynamics, θ , concerns the description of a simple diffusion operator ($\partial/\partial t = -(\delta Y^6/\theta) \times \Delta^3$). Here, the operator Δ^3 is an iterated laplacian operator equivalent to $\Delta(\Delta(\Delta))$. θ is the dissipation timescale for structures of typical length δY (model grid) and the quantity $K = \delta Y^6/\theta$ is the diffusion coefficient. This operator only acts on the latitudinal profiles of horizontal winds u and v , the zonal and latitudinal wind components respectively. The dynamical dissipation timescale θ is set to 500 s in the stratosphere. With this value, our modelled prediction of u matches the observed zonal wind⁹. Then, for structures of typical length L the timescale increases as L^6/K , which preferentially dissipates the small structures. We have reduced the aerosol's distribution grid to 10 bins instead of 45 bins in the original model²⁶. The volume ratio between adjacent bins is 16. This still adequately reproduces the growth path of aerosols.

Received 28 March; accepted 2 July 2002; doi:10.1038/nature00961.

1. Lellouch, E. *et al.* Titan's atmosphere and hypothesis ocean: a re-analysis of the Voyager 1 radio-occultation and IRIS 7.7- μ m data. *Icarus* **79**, 328–349 (1989).
2. Sromovsky, L. A. *et al.* Implication of Titan's north-south brightness asymmetry. *Nature* **292**, 698–702 (1981).
3. Hordun, F. *et al.* Numerical simulation of the general circulation of the atmosphere of Titan. *Icarus* **117**, 358–374 (1995).
4. Tokano, T., Neubauer, F. M., Laube, M. & McKay, C. P. Seasonal variation of Titan's atmospheric structure simulated by a general circulation model. *Icarus* **47**, 493–520 (1999).
5. Rages, K. & Pollack, J. B. Vertical distribution of scattering haze in Titan's upper atmosphere. *Icarus* **55**, 50–62 (1983).
6. Young, E. F., Rannou, P., McKay, C. P., Griffith, C. A. & Noll, K. A three dimensional map of Titan's tropospheric haze distribution based on HST imaging. *Astron. J.* **123**, 3473–3486 (2002).
7. Samuelson, R. E., Mayo, L. A., Knuckles, M. A. & Khanna, R. J. C_4N_2 ice in Titan's north polar stratosphere. *Planet. Space Sci.* **45**, 941–948 (1997).
8. Coustenis, A. *et al.* Images of Titan at 1.3 and 1.6 micron with adaptive optics at the CFHT. *Icarus* **154**, 501–515 (2002).
9. Hubbard, W. B. *et al.* The occultation of 28 Sgr by Titan. *Astron. Astrophys.* **269**, 541–563 (1993).
10. Kostiuik, T. *et al.* Direct measurement of winds on Titan. *Geophys. Res. Lett.* **28**, 2361–2364 (2001).
11. Flasar, F. H. & Conrath, B. J. Titan's stratospheric temperatures: A case for dynamical inertia? *Icarus* **85**, 346–654 (1990).
12. Coustenis, A. & Bézard, B. Titan's atmosphere from Voyager infrared observations: IV. Latitudinal variations of temperature and composition. *Icarus* **115**, 126–140 (1995).

13. Maguire, W. C., Hanel, R. A., Jennings, D. E., Kunde, V. G. & Samuelson, R. E. C_3H_8 and C_3H_4 in Titan's atmosphere. *Nature* **292**, 683–686 (1981).
14. Kunde, V. G. *et al.* C_4H_2 , HC_3N , and C_2N_2 in Titan's atmosphere. *Nature* **292**, 686–688 (1981).
15. Khare, B. N. *et al.* Optical constants of organic tholins produced in a simulated Titanian atmosphere: from soft X-ray to microwave frequencies. *Icarus* **60**, 127–137 (1984).
16. McKay, C. P., Pollack, J. B. & Courtin, R. The thermal structure of Titan's atmosphere. *Icarus* **80**, 23–53 (1989).
17. Danielson, R. E., Caldwell, J. J. & Larach, D. R. An inversion in the atmosphere of Titan. *Icarus* **20**, 437–443 (1973).
18. McKay, C. P., Pollack, J. B. & Courtin, R. The greenhouse and antgreenhouse effects on Titan. *Science* **253**, 1118–1123 (1991).
19. Hutzell, W. T., McKay, C. P. & Toon, O. B. Effects of the time-varying haze production on Titan's geometric albedo. *Icarus* **105**, 162–174 (1993).
20. Pollack, J. B., Rages, K., Toon, O. B. & Yung, Y. L. On the relationship between secular brightness changes of Titan and solar variability. *Geophys. Res. Lett.* **7**, 829–832 (1980).
21. Allen, M., Pinto, J. P. & Yung, Y. L. Titan: aerosol photochemistry and variations related to sunspot cycle. *Astrophys. J.* **242**, L125–L128 (1980).
22. Rannou, P., Cabane, M. & Chassefière, E. Growth of aerosols in Titan's atmosphere and related time scales: A stochastic approach. *Geophys. Res. Lett.* **20**, 967–970 (1993).
23. Toon, O. B., McKay, C. P., Griffith, C. A. & Turco, R. P. A physical model of Titan aerosols. *Icarus* **95**, 24–53 (1992).
24. Hutzell, W. T., McKay, C. P., Toon, O. B. & Hourdin, F. Simulations of Titan's brightness by a two dimensional haze model. *Icarus* **119**, 112–129 (1995).
25. Cabane, M., Chassefière, E. & Israel, G. Formation and growth of photochemical aerosols in Titan's atmosphere. *Icarus* **96**, 176–189 (1992).
26. Cabane, M., Rannou, P., Chassefière, E. & Israel, G. Fractal aggregates in Titan's atmosphere. *Planet. Space Sci.* **41**, 257–267 (1993).
27. Lebonnois, S., Toubanc, D., Hourdin, F. & Rannou, P. Seasonal variations of Titan's atmospheric composition. *Icarus* **152**, 384–406 (2001).
28. Rannou, P., McKay, C. P., Botet, R. & Cabane, M. Semi-empirical model of absorption and scattering by isotropic fractal aggregates of spheres. *Planet. Space Sci.* **47**, 385–396 (1999).

Acknowledgements

This work was partially supported by the NASA Planetary Atmospheres Program and the French Programme National de Planétologie. P.R. thanks the National Research Council Associateship Program.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.R. (e-mail: pra@ccr.jussieu.fr).

Emergent excitations in a geometrically frustrated magnet

S.-H. Lee*, C. Broholm†, W. Ratcliff‡, G. Gasparovic†, Q. Huang*
T. H. Kim‡§ & S.-W. Cheong‡

* NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA

† Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA

‡ Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA

Frustrated systems are ubiquitous^{1–3}, and they are interesting because their behaviour is difficult to predict; frustration can lead to macroscopic degeneracies and qualitatively new states of matter. Magnetic systems offer good examples in the form of spin lattices, where all interactions between spins cannot be simultaneously satisfied⁴. Here we report how unusual composite spin degrees of freedom can emerge from frustrated magnetic interactions in the cubic spinel $ZnCr_2O_4$. Upon cooling, groups of six spins self-organize into weakly interacting antiferromagnetic loops, whose directors—the unique direction along which the spins are aligned, parallel or antiparallel—govern all low-tem-

perature dynamics. The experimental evidence comes from a measurement of the magnetic form factor by inelastic neutron scattering; the data show that neutrons scatter from hexagonal spin clusters rather than individual spins. The hexagon directors are, to a first approximation, decoupled from each other, and hence their reorientations embody the long-sought local zero energy modes for the pyrochlore lattice.

Magnetism in transition metal oxides stems from atomic spins on the vertices of a periodic lattice. In insulators, interactions generally favour antiparallel nearest-neighbour spin alignment. For a simple cubic lattice (Fig. 1a), only a long-range ordered spin configuration can satisfy all interactions. On cooling, such systems show a continuous increase in the spin correlation length, culminating in a phase transition to long-range order. But for spins on the vertices of corner-sharing tetrahedra (Fig. 1b), no configuration can satisfy all interactions—a magnetic predicament called ‘geometrical frustration’⁴. Because the spin interaction energy is minimized when the four spins on each tetrahedron add to zero, interactions do not call for a divergent correlation length, but simply define a restricted phase space for fluctuations, parametrized by θ and ϕ (Fig. 1b) for each tetrahedron^{5,6}. Just as composite fermions can emerge from degenerate Landau levels in a two-dimensional electron gas⁷, the near-degenerate manifold of states in a frustrated magnet is fertile ground for emergent behaviour⁸.

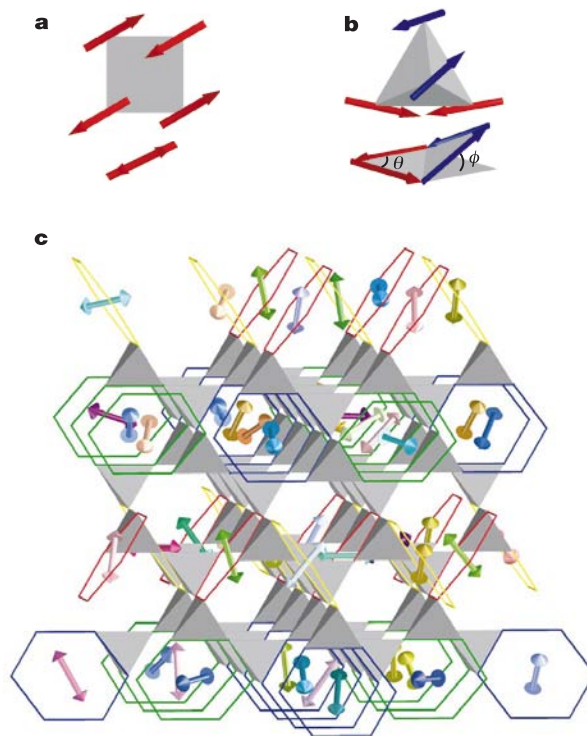


Figure 1 Lowest-energy spin configurations for four antiferromagnetically interacting spins on a square, a tetrahedron, and the pyrochlore lattice of corner-sharing tetrahedra. **a**, Setting aside global rotations, four spins on the vertices of a square with nearest-neighbour interactions have a unique lowest-energy spin configuration. **b**, For four spins on the vertices of a tetrahedron, any configuration with vanishing total spin has the lowest configuration energy. **c**, The lattice of corner-sharing tetrahedra formed by the octahedrally coordinated B sites in a spinel structure with chemical formula AB_2O_4 . A periodic assignment of all spins in the pyrochlore lattice is made to four different types of non-overlapping hexagons, represented by the colours blue, green, red and gold. Every spin belongs to just one hexagon, and each such hexagon carries a six spin director. The resulting tetragonal structure of these hexagons has a unit cell of $2a \times 2a \times 3c$, and can be described by a stacking of two different types of three-layer slabs along the c axis. The hexagon coverage on consecutive slabs is in fact uncorrelated, so that a macroscopic number of random slab-sequences can be generated.

§ Present address: Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Although direct structural probes are unavailable for quantum Hall samples, we can monitor the correlated spin state in a frustrated magnet by scattering neutrons from it⁹. Neutrons carry a magnetic dipole moment, and are subject to forces from atomic-scale field gradients. The resulting pattern of quasi-elastic scattering versus wavevector transfer, $\mathbf{Q} = \mathbf{k}_i - \mathbf{k}_f$, is the sample-averaged Fourier transform of spin configurations within a coherence volume of order $(100 \text{ \AA})^3$ given by the instrumental resolution. Here $\mathbf{k}_i$ and $\mathbf{k}_f$ are the de Broglie wavevectors associated with the incident and scattered neutrons, respectively. Magnetic peaks generally sharpen with decreasing temperature as the correlation length, ξ , increases. For un-frustrated La_2CuO_4 (ref. 10), the half-width at half-maximum (HWHM), $\kappa(T) = \xi(T)^{-1}$, becomes indistinguishable from zero below a microscopic energy scale, the Curie–Weiss temperature $|\Theta_{\text{CW}}|$. In contrast, for frustrated ZnCr_2O_4 (Fig. 2), κ remains finite even below $|\Theta_{\text{CW}}|$, and extrapolates to a finite value as $T/|\Theta_{\text{CW}}| \rightarrow 0$.

The finite low-temperature correlation length in ZnCr_2O_4 signals the emergence of confined nanometre-scale spin clusters. Rather than being associated with temperature-dependent short-range order above a phase transition, the wavevector dependence of the low-temperature intensity (Fig. 3a, b) can be interpreted as a spin cluster form factor. As opposed to the form factor for an atomic spin, the cluster form factor vanishes for $Q \rightarrow 0$, which is evidence that clusters carry no net spin¹¹. Further analysis is complicated by the anisotropy of spin clusters, which can occur in four different orientations for a cubic crystal. Rather than Fourier-inverting the data, we therefore compare them to the orientationally averaged Fourier transform of potential spin clusters. Individual tetrahedra would be prime candidates, as they constitute the basic motif of the pyrochlore lattice. However, a tetrahedron is too small to account for the features observed (Fig. 3a, b). The next-smallest symmetric structural unit is the hexagonal loop formed by a cluster of six tetrahedra (Fig. 4a). Two spins from each tetrahedron occupy the vertices of a hexagon, while the other two spins from each tetrahedron belong to different hexagons. Averaging over the four different hexagon orientations in the pyrochlore lattice, the square of the antiferromagnetic hexagon spin-loop form factor can be

written as

$$|F_6(\mathbf{Q})|^2 \propto \left\{ \sin \frac{\pi}{2} h \cdot \left(\cos \frac{\pi}{2} k - \cos \frac{\pi}{2} l \right) \right\}^2 + \left\{ \sin \frac{\pi}{2} k \cdot \left(\cos \frac{\pi}{2} l - \cos \frac{\pi}{2} h \right) \right\}^2 + \left\{ \sin \frac{\pi}{2} l \cdot \left(\cos \frac{\pi}{2} h - \cos \frac{\pi}{2} k \right) \right\}^2$$

The magnetic neutron scattering intensity would follow $I_0(\mathbf{Q}) = |F_6(\mathbf{Q})|^2 |F(\mathbf{Q})|^2$, where $F(\mathbf{Q})$ is the magnetic form factor for Cr^{3+} . The excellent qualitative agreement between model and data in Fig. 3 provides compelling evidence that neutrons scatter from antiferromagnetic hexagonal spin clusters rather than individual spins. In effect, ZnCr_2O_4 at low temperatures is not a system of strongly interacting spins, but a ‘protectorate’ of weakly interacting spin-loop directors. (The term ‘protectorate’ was introduced⁸ to describe stable states of matter in strongly correlated many-body systems. As antiferromagnetic hexagonal spin loops appear to be stable composite degrees of freedom for the pyrochlore lattice, we call the corresponding low-temperature state of the frustrated magnet a protectorate.)

Thermal and quantum fluctuations that violate collinearity within the hexagons should induce residual interactions between directors. Such interactions may account for the inelasticity of the scattering, the director correlations reflected in the greater sharpness of the experimental features in Figs 2 and 3, and the increase of κ with T , which indicates gradual disintegration of the directors.

What is the basis for the emergence of spin-loop directors as the effective degrees of freedom in this frustrated magnet? Fig. 4 shows the spins surrounding a hexagon in the pyrochlore lattice. Spin configurations that satisfy all interactions are characterized by the connected vectors of Fig. 4b. Although the spin configuration

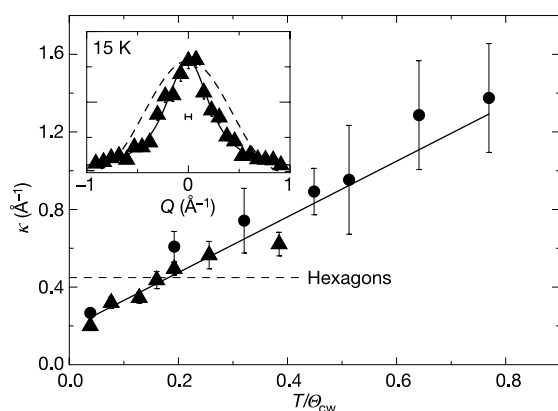


Figure 2 Temperature dependence of the inverse correlation length, $\kappa(T) = \xi(T)^{-1}$. (Temperature is given in units of the Curie–Weiss temperature, Θ_{CW} .) The data were derived from antiferromagnetic neutron scattering peaks by fitting to resolution-convoluted lorentzians. The triangles and circles are the lorentzian HWHMs along the $(h/4 \ 5/4)$ direction for $\hbar\omega = 1 \text{ meV}$, obtained with seven and eleven analyser blades, respectively. κ does not vanish as $T/|\Theta_{\text{CW}}| \rightarrow 0$, but extrapolates to a value that is close to the HWHM associated with the squared form factor for antiferromagnetic hexagon spin loops (dashed line). Inset, raw data for ZnCr_2O_4 at $T = 15 \text{ K}$. The bar shows the experimental resolution. The solid line is a resolution convoluted two-dimensional lorentzian; the dashed line is the squared hexagon spin-loop form factor convoluted with resolution.

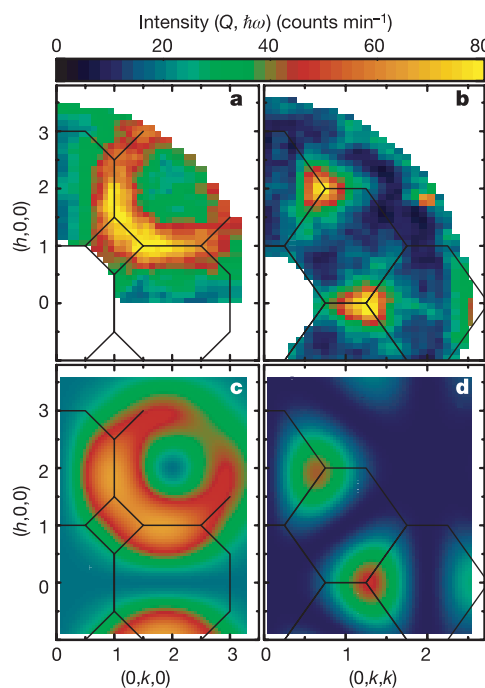


Figure 3 Wavevector dependence of the inelastic neutron scattering cross-section for ZnCr_2O_4 . **a, b**, Colour images of inelastic neutron scattering intensities from single crystals of ZnCr_2O_4 in the $(hk0)$ and (hkk) symmetry planes obtained at $T = 15 \text{ K}$ for $\hbar\omega = 1 \text{ meV}$. The data are a measure of the dynamic form factor for self-organized nanometre-scale spin clusters in the material. **c, d**, Colour images of the form factor squared calculated for antiferromagnetic hexagon spin loops averaged over the four hexagon orientations in the spinel lattice. The excellent agreement between model and data identifies the spin clusters as hexagonal spin loops.

remains severely under-constrained, no adjustment is possible without affecting spins on the outer perimeter. A general lowest-energy spin configuration on the pyrochlore lattice should therefore have wave-like soft modes that extend throughout the lattice. However, if the six hexagon spins are antiparallel with each other (Fig. 4c), then the staggered magnetization vector for a single hexagon, which for brevity we shall call the spin-loop director, is decoupled from the 12 outer spins, and hence its reorientation embodies a local zero energy mode for the pyrochlore lattice¹². It is possible to assign simultaneously all spins on the spinel lattice to hexagons, thus producing $N/6$ (where N is total number of spins) weakly interacting degrees of freedom (Fig. 1c). Accordingly, states that account for $1/6$ of the entropy of the magnet are accessible through local fluctuations from configurations where all spins are bunched into directors. Indeed, the measured entropy of $0.15 R \ln 4$ (where R is the gas constant) just above a low-temperature spin-Peierls-like phase transition^{4,13} is close to the predicted entropy of $(1/6) R \ln 4$ for uncorrelated directors in a spin-3/2 magnet. In contrast, there are no local soft modes for a general spin configuration in the low-energy manifold. This distinction between a general state from the low-energy manifold and states from the director protectorate is probably the basis for the stability of the latter.

Our finding that fluctuations in ZnCr_2O_4 involve spin clusters and not individual spins provides a natural explanation for a range of properties of geometrically frustrated magnets. It explains why the temperature-dependent susceptibility of frustrated magnets is accurately described by exact diagonalization of judiciously chosen spin clusters^{14–16}. Moreover, the so-called ‘undecouplable’ muon spin resonance response¹⁷ is recognized as being a consequence of muons sensing slow, large-amplitude fluctuations of select spins in a spin director. A director protectorate also provides a natural explanation for the coexistence of low characteristic energy scales with rigid short-range order, as evidenced by specific-heat¹⁸ and quasi-elastic neutron data¹⁹.

The director protectorate hints at an organizing principle for frustrated systems: if macroscopic condensation is not possible, interacting degrees of freedom combine to form rigid composite entities or clusters with weak mutual interactions. Exploring the

generality and basis for such a principle should be an interesting focus for theoretical work and experiments on a wider class of systems. Composite degrees of freedom are common in strongly interacting many-body systems. Quarks form hadrons, hadrons form nuclei, nuclei plus electrons form atoms, and atoms form molecules, which in turn are the basis for complex biological functionality. Planets, stars, galaxies and galaxy clusters are examples of clustering on a grander length scale. However, to our knowledge, the emergence of a confined spin cluster degree of freedom has not previously been documented in a uniform gapless magnet. The discovery is important because magnets offer an opportunity not afforded by the above-mentioned systems to monitor emergent structure in complex interacting systems with microscopic probes such as neutron scattering and NMR. The collapse of a geometrically frustrated magnet into a director protectorate could, for example, be a useful template for exploring aspects of protein folding^{2,20}. □

Methods

Three crystals of ZnCr_2O_4 (total mass 200 mg) were co-mounted for the inelastic neutron scattering measurements. The measurements were performed using the cold neutron triple-axis spectrometer SPINS at the National Institute of Standards and Technology Center for Neutron Research. A vertically focusing pyrolytic graphite (002) monochromator, PG(002), extracted a monochromatic beam with energy $E_i = 6.1$ meV from a ^58Ni -coated cold neutron guide. Scattered neutrons were analysed with seven or eleven $2.1 \text{ cm} \times 15 \text{ cm}$ PG(002) analyser blades that reflected neutrons with $E_f = 5.1$ meV onto a ^3He proportional counter. Cooled Be filtered the scattered beam. Received 22 February; accepted 4 July 2002; doi:10.1038/nature00964.

1. Debenetti, P. G. & Stillinger, F. H. Supercooled liquids and the glass transition. *Nature* **410**, 259–267 (2001).
2. Wolynes, P. G. & Eaton, W. A. The physics of protein folding. *Phys. World* **12**, 39–44 (1999).
3. Bramwell, S. T. & Gingras, M. J. P. Spin ice state in frustrated magnetic pyrochlore materials. *Science* **294**, 1495–1501 (2001).
4. Ramirez, A. P. in *Handbook on Magnetic Materials* (ed. Busch, K. J. H.) Vol. **13**, 423–520 (Elsevier Science, Amsterdam, 2001).
5. Moessner, R. & Chalker, J. T. Properties of a classical spin liquid: the Heisenberg pyrochlore antiferromagnet. *Phys. Rev. Lett.* **80**, 2929–2932 (1998).
6. Canals, B. & Lacroix, C. Pyrochlore antiferromagnet: a three-dimensional quantum spin liquid. *Phys. Rev. Lett.* **80**, 2933–2936 (1998).
7. Stormer, H. L., Tsui, D. C. & Gossard, A. C. The fractional quantum Hall effect. *Rev. Mod. Phys.* **71**, S298–S305 (1999).
8. Laughlin, R. B. & Pines, D. The theory of everything. *Proc. Natl Acad. Sci. USA* **97**, 28–31 (2000).
9. Lovesey, S. W. *Theory of Thermal Neutron Scattering from Condensed Matter* (Clarendon, Oxford, 1984).
10. Birgeneau, R. J. et al. Instantaneous spin correlations in La_2CuO_4 . *Phys. Rev. B* **59**, 13788–13794 (1999).
11. Lee, S.-H. et al. Spin-glass and non-spin-glass features of a geometrically frustrated magnet. *Europhys. Lett.* **35**, 127–132 (1996).
12. Moessner, R. & Chalker, J. T. Low-temperature properties of classical, geometrically frustrated antiferromagnets. *Phys. Rev. B* **58**, 12049–12062 (1998).
13. Lee, S.-H., Broholm, C., Kim, T. H., Rattcliff, W. & Cheong, S.-W. Local spin resonance and spin-Peierls-like phase transition in a geometrically frustrated antiferromagnet. *Phys. Rev. Lett.* **84**, 3718–3721 (2000).
14. Moessner, R. & Berlinsky, A. J. Magnetic susceptibility of diluted pyrochlore and $\text{SrCr}_{9-x}\text{Ga}_{3+x}\text{O}_{19}$. *Phys. Rev. Lett.* **83**, 3293–3296 (1999).
15. Garcia-Adeva, A. J. & Huber, D. L. Quantum tetrahedral mean-field theory of the pyrochlore lattice. *Can. J. Phys.* **79**, 1359–1364 (2001).
16. Mekata, M. & Yamada, Y. Magnetic-ordering process in $\text{SrCr}_2(\text{Ga-In})_2\text{O}_{19}$. *Can. J. Phys.* **79**, 1421–1426 (2001).
17. Uemura, Y. J. et al. Spin fluctuations in frustrated kagomé lattice system $\text{SrCr}_8\text{Ga}_4\text{O}_{19}$ studied by muon spin relaxation. *Phys. Rev. Lett.* **73**, 3306–3309 (1994).
18. Ramirez, A. P., Espinosa, G. P. & Cooper, A. S. Elementary excitations in a diluted antiferromagnetic kagomé lattice. *Phys. Rev. B* **45**, 2505–2508 (1992).
19. Broholm, C., Aeppli, G., Espinosa, G. P. & Cooper, A. S. Antiferromagnetic fluctuations and short range order in a kagomé lattice. *Phys. Rev. Lett.* **65**, 3173–3176 (1990).
20. Pande, V. S., Grosberg, A. Yu. & Tanaka, T. Heteropolymer freezing and design: Towards physical models of protein folding. *Rev. Mod. Phys.* **72**, 259–314 (2000).

Acknowledgements

We thank O. Tchernyshyov, R. Moessner, S. L. Sondhi, A. B. Harris, G. Aeppli, N. Read and D. Weitz for discussions, J. J. Rush, A. P. Ramirez and P. M. Gehring for critical reading of the manuscript, and Z. Huang for assistance in making figures. This work was partially supported by the NSF and the BSF.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.H.L. (e-mail: shl@nist.gov).

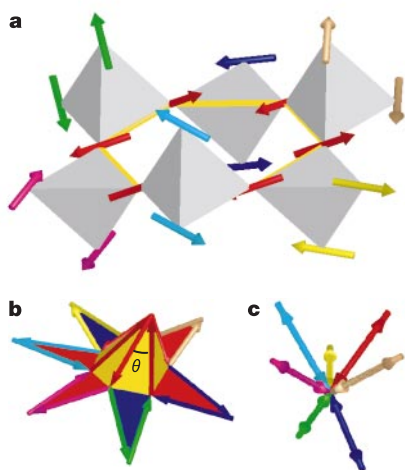


Figure 4 Possible spin fluctuations in the classical ground-state manifold. **a**, Spin cluster surrounding a hexagon (shown in yellow) in the pyrochlore lattice of Fig. 1c. **b**, Generalized pattern of classical spin vectors on six neighbouring tetrahedra satisfying the condition that there be no net spin on any tetrahedron. **c**, Pattern of spin vectors satisfying the condition that $\theta = 0$ for all tetrahedra. For such spin configurations, each tetrahedron has two pairs of antiparallel spins, and each hexagon has six collinear and antiferromagnetic spins as indicated by six red arrows in **a**. The energy of the spin cluster is independent of the orientation of the spin-loop directors.

Chemical investigation of hassium (element 108)

Ch. E. Düllmann^{*†}, W. Bröchle[‡], R. Dressler[†], K. Eberhardt[§], B. Eichler[†], R. Eichler[†], H. W. Gäggeler^{*†}, T. N. Ginter^{||}, F. Glaus[†], K. E. Gregorich^{||}, D. C. Hoffman[¶], E. Jäger[‡], D. T. Jost[†], U. W. Kirbach^{||}, D. M. Lee^{||}, H. Nitsche[¶], J. B. Patin[¶], V. Pershina[‡], D. Piguet[†], Z. Qin[#], M. Schädel[‡], B. Schausten[‡], E. Schimpf[‡], H.-J. Schött[‡], S. Soverna^{*†}, R. Sudowe^{||}, P. Thörle[§], S. N. Timokhin[☆], N. Trautmann[§], A. Türler^{*†}, A. Vahle^{††}, G. Wirth[‡], A. B. Yakushev[☆] & P. M. Zielinski^{||}

^{*} Departement für Chemie und Biochemie, Universität Bern, CH-3012 Bern, Switzerland

[†] Labor für Radio- und Umweltchemie, Paul Scherrer Institut, CH-5232 Villigen, Switzerland

[‡] Gesellschaft für Schwerionenforschung mbH, D-64291 Darmstadt, Germany

[§] Institut für Kernchemie, Universität Mainz, D-55128 Mainz, Germany

^{||} Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

[¶] Department of Chemistry, University of California, Berkeley, California 94720-1460, USA

[#] Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, P.R. China

[☆] Flerov Laboratory of Nuclear Reactions, Joint Institute for Nuclear Research, 141980 Dubna, Russia

^{**} Institut für Radiochemie, Technische Universität München, D-85748 Garching, Germany

^{††} Research Center Rossendorf e.V., D-01314 Dresden, Germany

The periodic table provides a classification of the chemical properties of the elements. But for the heaviest elements, the transactinides, this role of the periodic table reaches its limits because increasingly strong relativistic effects on the valence electron shells can induce deviations from known trends in chemical properties^{1–4}. In the case of the first two transactinides, elements 104 and 105, relativistic effects do indeed influence their chemical properties⁵, whereas elements 106 and 107 both behave as expected from their position within the periodic

table^{6,7}. Here we report the chemical separation and characterization of only seven detected atoms of element 108 (hassium, Hs), which were generated as isotopes ²⁶⁹Hs (refs 8, 9) and ²⁷⁰Hs (ref. 10) in the fusion reaction between ²⁶Mg and ²⁴⁸Cm. The hassium atoms are immediately oxidized to a highly volatile oxide, presumably HsO₄, for which we determine an enthalpy of adsorption on our detector surface that is comparable to the adsorption enthalpy determined under identical conditions for the osmium oxide OsO₄. These results provide evidence that the chemical properties of hassium and its lighter homologue osmium are similar, thus confirming that hassium exhibits properties as expected from its position in group 8 of the periodic table.

The discovery of Hs was reported in 1984 (ref. 11) with the identification of the nuclide ²⁶⁵Hs with a half-life of $T_{1/2} = 1.5$ ms (refs 11, 12). In 1996, the much longer-lived isotope ²⁶⁹Hs, with a half-life of about 10 s, was observed in the α -decay chain of ²⁷⁷112 (ref. 8). Recently, evidence for the existence of the neighbouring nuclide ²⁷⁰Hs was found and a half-life of about 4 s was deduced from its measured α -decay energy¹⁰. The latter two Hs isotopes might be sufficiently long-lived to allow their chemical characterization. Suitable reactions for their direct production are ²⁴⁸Cm(²⁶Mg, 5,4n)^{269,270}Hs, for which formation cross-sections of a few picobarn have been estimated¹³.

The periodic table suggests that Hs is a member of group 8 and thus chemically similar to its lighter homologues ruthenium (Ru) and osmium (Os), which are known to form highly volatile tetroxides. Therefore, Hs is also expected to form a very volatile tetroxide (HsO₄) suitable for gas-phase isolation^{1,14–16}, even though two earlier attempts to chemically identify Hs in the tetroxide form proved unsuccessful^{17,18}.

Fully relativistic density functional calculations¹⁹ for the tetroxides of the group 8 elements indicated that the electronic structure of HsO₄ is very similar to that of OsO₄, with covalent bonding being somewhat more pronounced in the former. The stability of the gaseous tetroxides was found¹⁹ to increase in the order RuO₄ < OsO₄ < HsO₄, in agreement with extrapolations within group 8 of the periodic table²⁰. The density functional calculations, in conjunction with a surface interaction model, suggest adsorption

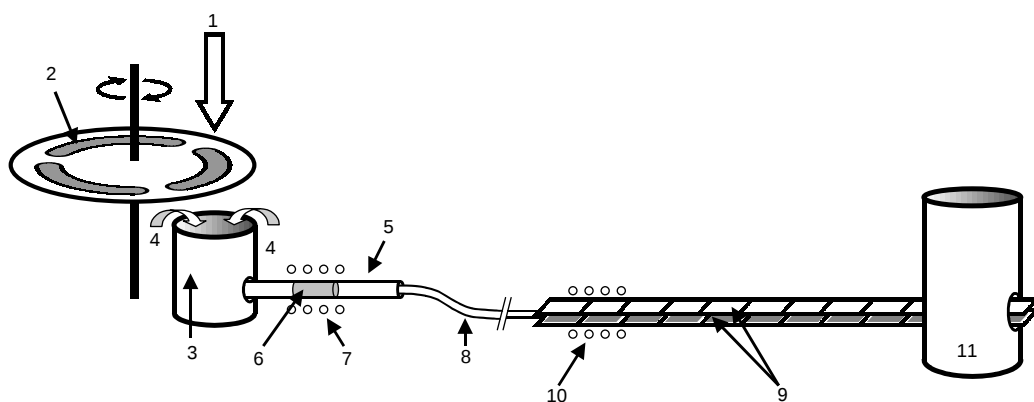


Figure 1 Schematic drawing of the IVO-COLD set-up used to produce and isolate Hs isotopes in form of the volatile HsO₄. The ²⁶Mg-beam (1) passes through the rotating vacuum window and ²⁴⁸Cm-target (2) assembly. The target consisted of three banana-shaped segments (1.9 cm² area each) covered with 239 μ g cm⁻², 730 μ g cm⁻², and 692 μ g cm⁻² ²⁴⁸Cm, respectively. The ²⁴⁸Cm (isotopic composition ²⁴⁶Cm: 4.2%; ²⁴⁸Cm: 95.8%) was deposited on 2.82 mg cm⁻² beryllium (Be) backings by molecular plating. The target-window assembly rotated in the adjacent gas volume with 2,000 rev min⁻¹ and was synchronized with the beam macrostructure of the accelerator in order to distribute each 6-ms beam pulse evenly over one target segment. In the fusion reaction ^{269,270}Hs nuclei are formed that recoil out of the target into a gas volume (3) and are flushed with a He/O₂ mixture (4) out of the chamber. The gas was passed through a

cartridge containing P₂O₅ as a drying agent before injecting it into IVO. In this way, the water vapour concentration was reduced to a measured value of <1 p.p.m. throughout the experiment. The gas was injected into a quartz column (5) containing a quartz wool plug (6) heated to 600 °C by an oven (7). There, Hs is converted to HsO₄ which is volatile at room temperature and transported with the gas flow through a 10-m-long perfluoroalkoxy (PFA) Teflon capillary (8) to the COLD detector array registering the nuclear decay (α and spontaneous fission) of the Hs nuclides. The array consists of 24 detectors arranged in 12 pairs (9). A temperature gradient was established along the detector array by means of a thermostat (10) at the entrance and a liquid nitrogen cryostat (11) at the exit. The temperature was monitored by five thermocouples installed along the copper bar. Depending on the volatility of HsO₄, the molecules adsorb at a characteristic temperature.

enthalpies of HsO_4 and OsO_4 on a quartz surface of $(-35.9 \pm 1.5) \text{ kJ mol}^{-1}$ and $(-38.0 \pm 1.5) \text{ kJ mol}^{-1}$, respectively. Extrapolations of the volatility of group 8 tetroxides suggest almost identical adsorption enthalpies of $(-46 \pm 15) \text{ kJ mol}^{-1}$ and $(-45 \pm 15) \text{ kJ mol}^{-1}$ for HsO_4 and OsO_4 , respectively, whereas a physisorption model yields identical values of $(-47 \pm 11) \text{ kJ mol}^{-1}$ for both tetroxides. Overall, the theoretical values suggest fairly similar adsorption behaviour for HsO_4 and OsO_4 on quartz. Although the present experiment uses detectors covered with a silicon nitride layer, we expect a direct analogy between theory and experiment because we observed comparable adsorption interactions of OsO_4 with quartz and silicon nitride.

The expected high volatility of group 8 tetroxides allows excellent separation from heavy actinides and lighter transactinides as well as from Pb, Bi and Po, which is important because several isotopes of these elements are formed with high yield as byproducts of the nuclear fusion reaction. These nuclides often decay via emission of α -particles and, therefore, severely interfere with the unambiguous detection of the nuclear decay of the investigated Hs nuclide.

Online investigations of Os oxides have already been conducted earlier^{15,21,22}. Using the *in situ* volatilization and online detection apparatus (IVO) carrier-free $^{173}\text{OsO}_4$ ($T_{1/2} = 22.4 \text{ s}$) could be separated and detected with an overall efficiency of $(40 \pm 10)\%$. Decontamination from Po ($\geq 2 \times 10^4$) was excellent²¹.

The experimental set-up, schematically shown in Fig. 1, involves mounting the ^{248}Cm target²³ on a rotating wheel and bombarding it with up to $8 \times 10^{12} \text{ }^{26}\text{Mg}^{5+}$ particles per second, delivered by the UNILAC accelerator at the Gesellschaft für Schwerionenforschung mbH. The particle beam first passed through a rotating three-segment 3.68 mg cm^{-2} Be vacuum window which allowed for a pressure up to 1.3 atm in the recoil chamber. The 192.7-MeV beam energy delivered by the UNILAC resulted in ^{26}Mg projectile energies of 143.7–146.8 MeV inside the ^{248}Cm target. Nuclear reaction

products recoiling from the target were thermalized in the gas volume (34 ml) of the IVO device²¹ flushed with dry (measured water vapour concentration $<1 \text{ p.p.m.}$) 1.2 l min^{-1} helium (He) and 100 ml min^{-1} oxygen (O_2). The reaction products were transported with the carrier gas through a 30-cm-long quartz column (internal diameter, i.d., 4 mm) containing a quartz wool plug at a distance of 6.5 cm from the recoil chamber. This plug was heated to 600°C and served as a filter for aerosol particles and provided a surface to complete the oxidation reaction of Os and Hs to their tetroxides, which were further transported through a 10-m-long perfluoroalkoxy (PFA) Teflon capillary (i.d., 2 mm) to the detection system. The capillary was installed inside a polyamide (PA) tube flushed with dry N_2 .

Using gas phase adsorption thermochromatography²⁴, we measured the temperature at which HsO_4 deposits, with a chromatographic column along which a stationary negative temperature gradient is maintained. The chromatographic column, the cryo online detector (COLD), also served as detection system for the identification of decaying atoms of $^{269,270}\text{Hs}$. COLD consists of 12 pairs of silicon PIN-photodiodes of $1 \times 3 \text{ cm}^2$ active area mounted at a distance of 1.5 mm via two spacers made from silicon. PIN diodes are well suited for detection of α -particles or fission fragments. The diode pairs were installed inside a copper bar. A temperature gradient from -20°C to -170°C was established along the detector array. The efficiency for detecting a single α -particle emitted by a species adsorbed within the detector array was 77%. To avoid formation of ice layers on the detector surfaces, the whole device was placed inside a vacuum tight housing flushed with dry N_2 . COLD is an improved version of a previous set-up called the cryo-thermochromatography separator (CTS)²².

The detectors of the COLD array were online calibrated with α -decaying ^{219}Rn and its daughters ^{215}Po and ^{211}Bi using a ^{227}Ac source. The determined energy resolution was 50–70 keV (full-

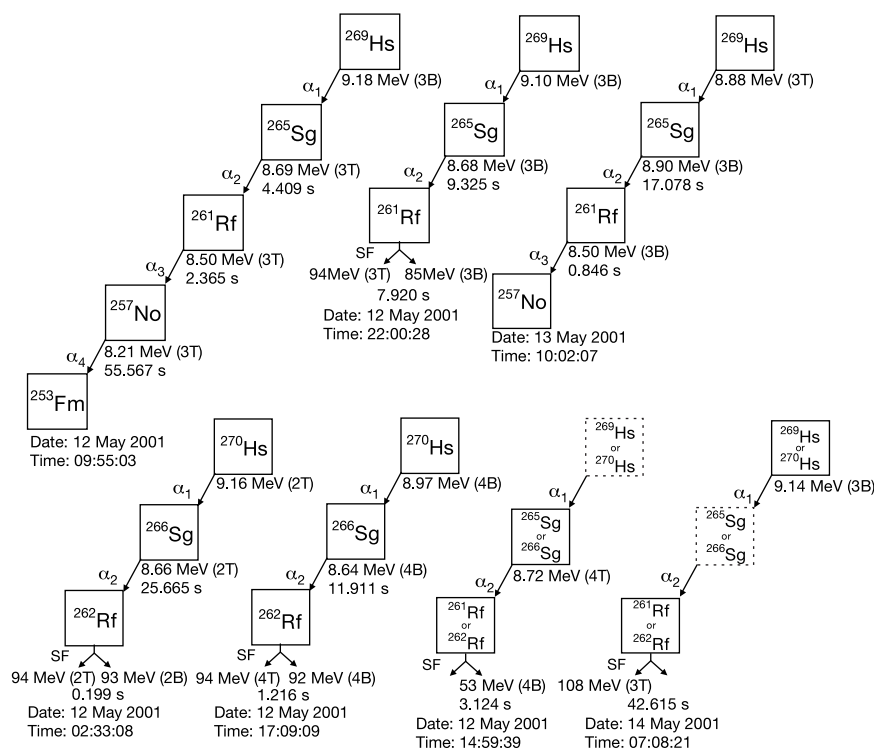


Figure 2 The seven nuclear decay chains originating from Hs isotopes that were detected in the course of the experiment permit an unambiguous identification of hassium after chemical separation. Indicated are the energies of α -particles and fission fragments in megaelectronvolts (MeV) and the lifetimes in seconds (s). The detector in which the decay

was registered is indicated in parentheses where T stands for top detector and B for bottom detector. For each chain the date and time of its registration are given. The lifetime of the mother isotope could not be determined with the applied thermochromatography technique because the deposition time is not measured.

width at half-maximum, FWHM) for detectors 1 through 8 and 80–110 keV (FWHM) for detectors 9 through 12.

The proper functioning of the IVO-COLD system was checked at the beginning and after the end of the Hs experiment by mounting a $800 \mu\text{g cm}^{-2} {}^{152}\text{Gd}$ (32% enriched) target to produce short-lived Os isotopes in the reaction ${}^{152}\text{Gd}({}^{26}\text{Mg}, 6n){}^{172}\text{Os}$. The beam intensity was $1 \times 10^{12} \text{ s}^{-1}$ and the beam energy in the middle of the target 153 MeV. All other experimental parameters, including gas composition and flow rate, were identical to those of the Hs experiment. In the COLD detector array the α -decay of ${}^{172}\text{Os}$ was registered. ${}^{172}\text{Os}$ has a half-life of 19.2 s and a 1.0% α -decay branch with $E_{\alpha} = 5.10 \text{ MeV}$.

The experiment to produce Hs isotopes lasted 64.2 h during which a total of $1.0 \times 10^{18} {}^{26}\text{Mg}$ particles passed through the target. Only α -lines originating from ${}^{211}\text{At}$, ${}^{219,220}\text{Rn}$ and their decay products were identified. Whereas ${}^{211}\text{At}$ and its decay product ${}^{211}\text{Po}$ were deposited mainly in the first two detectors, ${}^{219,220}\text{Rn}$ and their decay products accumulated in the last three detectors, where the temperature was sufficiently low to partly adsorb Rn. One side of detector pair 1 did not operate owing to a technical failure and, therefore, this detector unit was excluded from the data analysis. During the experiment, seven correlated decay chains were detected (see Fig. 2). All correlated α -decay chains were observed in detectors 2 through 4 and assigned to the decay of ${}^{269}\text{Hs}$ or ${}^{270}\text{Hs}$. The characteristics of the first three decay chains agree well with data^{8,9} on ${}^{269}\text{Hs}$ and its daughter nuclides, while two other decay chains can be attributed¹⁰ to the decay of ${}^{270}\text{Hs}$. The last two decay chains were incomplete and a definite assignment to ${}^{269}\text{Hs}$ or ${}^{270}\text{Hs}$ could not be made. No additional three-member decay chains within $\leq 300 \text{ s}$ were registered in detectors 2 to 10. The background count-rate of α -particles with energies between 8.0 and 9.5 MeV was about 0.6 h^{-1} per detector, leading to very low probabilities of $\leq 7 \times 10^{-5}$ and $\leq 2 \times 10^{-3}$ for any of the first five chains and any of the last two chains, respectively, being of random origin. In addition, four uncorrelated fission fragments with ener-

gies $> 50 \text{ MeV}$ were registered in detectors 2 through 4. All other detectors registered no fission fragments, except for one fission fragment being observed in the operating side of detector 1.

As depicted in Fig. 3, the α -decay of one Hs atom was registered in detector 2, the decay of four atoms in detector 3 and the decay of two atoms in detector 4. The maximum of the Hs distribution was found for a temperature of $(-44 \pm 6) ^\circ\text{C}$. The deposition distribution of OsO_4 measured before and after the experiment revealed a maximum in detector 6 at $(-82 \pm 7) ^\circ\text{C}$.

The adsorption enthalpy (ΔH_{ads}) of the compound on the stationary phase is extracted from the measured deposition distribution by using Monte Carlo simulations of the trajectories of single molecules as they move along the column under real experimental conditions²⁵. The only free parameter in the simulations is ΔH_{ads} , but the half-life of the nuclide is a crucial parameter. For this reason, and because the half-life of ${}^{270}\text{Hs}$ has not yet been measured, only decays assigned to ${}^{269}\text{Hs}$ were used to evaluate the adsorption enthalpy of the compound on the silicon nitride surface. The results that best reproduce the experimental data are shown in Fig. 3 (solid lines) and suggest a value of $\Delta H_{\text{ads}} = (-46 \pm 2) \text{ kJ mol}^{-1}$ (68% confidence interval, c.i.), which was inferred using a $T_{1/2}$ value of 11^{+15}_{-4} s for ${}^{269}\text{Hs}$ (refs 8, 9). The given uncertainty is the total uncertainty; that is, it reflects the width of the measured deposition peak and the uncertainty in $T_{1/2}$. The adsorption enthalpy of OsO_4 on silicon nitride deduced from this experiment was $(-39 \pm 1) \text{ kJ mol}^{-1}$, which is in good agreement with the adsorption enthalpy obtained in earlier investigations using quartz surfaces^{15,21,22}.

The experimentally derived adsorption enthalpy of HsO_4 is thus lower than that of OsO_4 , while the predictions suggest either similar values²⁰ or a slightly higher value¹⁹ for HsO_4 . However, the theoretical values and experimental values have associated uncertainties. Moreover, the low value of the ΔH_{ads} determined for the hassium oxide species clearly suggests that it is HsO_4 , given that, by analogy with the known properties of the Os oxides, other Hs oxides are all expected to be less volatile and unable to reach the detection system. The observed formation of a very volatile Hs molecule, presumably HsO_4 , in a mixture of oxygen and helium thus provides strong qualitative evidence that Hs is an ordinary member of group 8 of the periodic table that behaves similarly to its lighter homologue Os. □

Received 8 April; accepted 9 July 2002; doi:10.1038/nature00980.

1. Fricke, B. Superheavy elements. *Struct. Bonding* **21**, 90–144 (1975).
2. Pyykkö, P. & Desclaux, J.-P. Relativity and the periodic system of elements. *Acc. Chem. Res.* **12**, 276–281 (1979).
3. Pershina, V. G. Electronic structure and properties of the transactinides and their compounds. *Chem. Rev.* **96**, 1977–2010 (1996).
4. Schwerdtfeger, P. & Seth, M. Relativistic effects of the superheavy elements. *Encyclopedia of Computational Chemistry* Vol. 4 2480–2499 (Wiley, New York, 1998).
5. Kratz, J. V. in *Heavy Elements and Related New Phenomena* Ch. 4 (eds Greiner, W. & Gupta, R. K.) 129–193 (World Scientific, Singapore, 1999).
6. Schädel, M. *et al.* Chemical properties of element 106 (seaborgium). *Nature* **388**, 55–57 (1997).
7. Eichler, R. *et al.* Chemical characterization of bohrium (element 107). *Nature* **407**, 63–65 (2000).
8. Hofmann, S. *et al.* The new element 112. *Z. Phys. A* **354**, 229–230 (1996).
9. Hofmann, S. & Münzenberg, G. The discovery of the heaviest elements. *Rev. Mod. Phys.* **72**, 733–767 (2000).
10. Türler, A. *et al.* Decay properties of ${}^{269}\text{Hs}$ and evidence for the new nuclide ${}^{270}\text{Hs}$. *Eur. Phys. J. A* (submitted).
11. Münzenberg, G. *et al.* The identification of element 108. *Z. Phys. A* **317**, 235–236 (1984).
12. Hofmann, S. *et al.* Production and decay of ${}^{269}\text{Hs}$. *Z. Phys. A* **350**, 277–280 (1995).
13. Schädel, M. & Hofmann, S. Prospects for the discovery of new elements. *J. Radioanal. Nucl. Chem.* **203**, 283–300 (1996).
14. Bächmann, K. & Hoffmann, P. Chemische Probleme bei der Darstellung überschwerer Elemente durch Kernreaktionen. *Radiochim. Acta* **15**, 153–163 (1971).
15. Domanov, V. P. & Zvara, I. Continuous-flow thermochromatographic separation of unsupported radioisotopes of platinum elements in a stream of air from nuclear reaction products in an accelerator heavy-ion beam. *Sov. Radiochem.* **26**, 731–739 (1985); translated from *Radiokhimiya* **26**, 770–778 (1984).
16. Zude, F., Fan, W., Trautmann, N., Herrmann, G. & Eichler, B. Thermochromatography of platinum elements in oxygen: Radiochemical studies of the behaviour of rhodium, palladium, osmium and platinum. *Radiochim. Acta* **62**, 61–63 (1993).
17. Zhuikov, B. L., Kruz, H. & Zvara, I. Possibilities of chemical identification of short-lived isotopes of

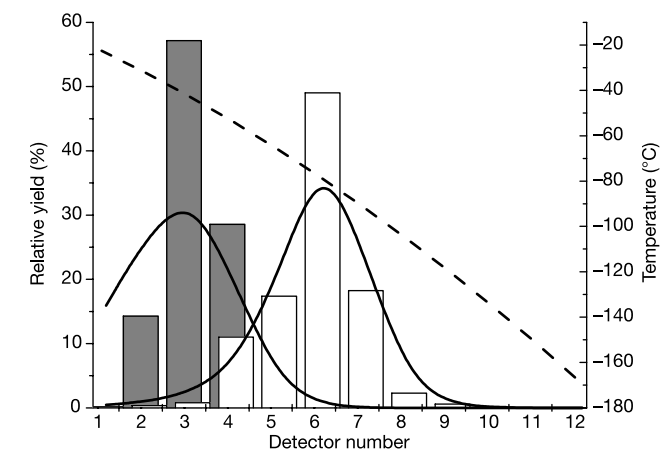


Figure 3 Merged thermochromatogram of HsO_4 and OsO_4 . Indicated are the relative yields of HsO_4 and OsO_4 for each of the 12 detector pairs. Measured values are represented by bars: HsO_4 , dark grey; OsO_4 , white. For Hs, the following distribution was measured: ${}^{269}\text{Hs}$: three events (all in detector 3); ${}^{270}\text{Hs}$: two events (one in detector 2 and one in detector 4); Hs (isotope unknown): two events (one in detector 3 and one in detector 4). For Os, the distribution of 1×10^5 events of ${}^{172}\text{OsO}_4$ is given. The dashed line indicates the temperature profile (right-hand scale). The maxima of the deposition distributions were evaluated as $(-44 \pm 6) ^\circ\text{C}$ for HsO_4 and $(-82 \pm 7) ^\circ\text{C}$ for OsO_4 where the uncertainties indicate the temperature range covered by the detector which registered the maximum of the deposition distribution. Solid lines represent results of a Monte Carlo simulation of the migration process of the species along the column with standard adsorption enthalpies of $-46.0 \text{ kJ mol}^{-1}$ for ${}^{269}\text{HsO}_4$ and $-39.0 \text{ kJ mol}^{-1}$ for ${}^{172}\text{OsO}_4$.

- element 108. Report P7-86-322, page 26 (Joint Institute for Nuclear Research, JINR, Dubna, 1986) (in Russian).
18. Dougan, R. J., Moody, K. J., Hulet, E. K. & Bethune, G. R. OSCAR: An apparatus for on-line gas-phase separations. FY87 Annual Report UCAR 10062/87, 4–17 (Lawrence Livermore National Laboratory, LLNL, Nuclear Chemistry Division, Livermore, 1987).
 19. Pershina, V., Bastug, T., Fricke, B. & Varga, S. The electronic structure and properties of group 8 oxides MO_x , where $M = \text{Ru, Os}$, and element 108, Hs. *J. Chem. Phys.* **115**, 792–799 (2001).
 20. Düllmann, Ch. E., Eichler, B., Eichler, R., Gägger, H. W. & Türler, A. On the stability and volatility of group 8 tetroxides MO_4 ($M = \text{ruthenium, osmium, and hassium (Z = 108)}$). *J. Phys. Chem. B* **106**, 6679–6684 (2002).
 21. Düllmann, Ch. E. *et al.* IVO, a device for in situ volatilization and on-line detection of products from heavy ion reactions. *Nucl. Instrum. Meth. A* **479**, 631–639 (2002).
 22. Kirbach, U. W. *et al.* The cryo-thermochromatographic separator (CTS): A new rapid separation and α -detection system for on-line chemical studies of highly volatile osmium and hassium ($Z = 108$) tetroxides. *Nucl. Instrum. Meth. A* **484**, 587–594 (2002).
 23. Malmbeck, R. *et al.* Separation of ^{248}Cm from a ^{252}Cf neutron source for production of Cm targets. *Radiochim. Acta* **89**, 543–549 (2001).
 24. Zvara, I. Thermochemical method of separation of chemical elements in nuclear and radiochemistry. *Isotopenpraxis* **26**, 251–258 (1990).
 25. Zvara, I. Simulation of thermochemical processes by the Monte Carlo method. *Radiochim. Acta* **38**, 95–101 (1985).

Acknowledgements

We thank the staff of the Laboratory for Micro- and Nanotechnology at PSI for manufacturing the PIN-diode sandwiches for the COLD array and the staff of the GSI UNILAC for providing stable, highly intense beams of ^{26}Mg as well as the target laboratory for Be foils for the vacuum windows. Support from the European Commission Institute for Transuranium Elements, Karlsruhe, for long-term storage of ^{252}Cf and the chemical separation of ^{248}Cm is appreciated. These studies were supported in part by the Swiss National Science Foundation and the Chemical Sciences Division of the Office of Basic Energy Sciences, US Department of Energy.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.W.G. (e-mail: gaeggeler@iac.unibe.ch).

Mantle compensation of active metamorphic core complexes at Woodlark rift in Papua New Guinea

Geoffrey A. Abers*, Aaron Ferris*, Mitchell Craig†‡, Hugh Davies†, Arthur L. Lerner-Lam§, John C. Mutter§ & Brian Taylor||

* Department of Earth Sciences, Boston University, Boston, Massachusetts 02215, USA

† Department of Geology, University of Papua New Guinea, Port Moresby, Papua New Guinea

§ Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA

|| University of Hawaii, Manoa, Hawaii 96822, USA

In many highly extended rifts on the Earth, tectonic removal of the upper crust exhumes mid-crustal rocks, producing metamorphic core complexes. These structures allow the upper continental crust to accommodate tens of kilometres of extension¹, but it is not clear how the lower crust and underlying mantle respond. Also, despite removal of the upper crust, such core complexes remain both topographically high and in isostatic equilibrium. Because many core complexes in the western United States are underlain by a flat Moho discontinuity^{2,3}, it has been widely assumed that their elevation is supported by flow in the lower crust^{4–6} or by magmatic underplating⁷. These processes should decouple upper-crust extension from that in the mantle.

‡ Present address: Department of Geological Sciences, California State University, Hayward, California 94542, USA.

In contrast, here we present seismic observations of metamorphic core complexes of the western Woodlark rift that show the overall crust to be thinned beneath regions of greatest surface extension. These core complexes are actively being exhumed⁸ at a rate of 5–10 km Myr^{−1}, and the thinning of the underlying crust appears to be compensated by mantle rocks of anomalously low density, as indicated by low seismic velocities. We conclude that, at least in this case, the development of metamorphic core complexes and the accommodation of high extension is not purely a crustal phenomenon, but must involve mantle extension.

The Woodlark rift of Papua New Guinea (Fig. 1) is the site of some of the youngest, most recently uplifted metamorphic core complexes (MCCs) on the planet⁸, and the most rapidly extending continental crust⁹. Extension began producing new sea floor along strike by 5–6 Myr ago¹⁰, leading to MCC exhumation on Fergusson and Goodenough islands in the Pliocene epoch, as indicated by sediment composition in adjacent basins⁸. Most workers infer that extension is driven by plate forces such as slab pull at other margins of the Solomon Sea¹¹ or by gravitational collapse¹⁰, similar to some large-extension rifts elsewhere but contrasting with active rifts such as the east Africa rift^{12,13}. Magnetic lineations require 100–200 km extension across the D'Entrecasteaux MCCs¹⁰, at 20–35 mm yr^{−1}, much of which must be accommodated on the north-dipping detachment shear zones bounding them. The MCCs show rapid uplift, with footwall rocks having experienced conditions of 700–900 °C and 5–6 kbar as recently as 3–4 Myr ago, and 4–5 kbar at similar temperatures at 1.5–2 Myr ago^{14,15}. The MCCs seem to be at present exhuming, as indicated by geomorphic¹⁶ and low-temperature geochronological indicators¹⁴.

In 1999 and 2000, we placed 19 broadband seismographs at 11 sites across the western Woodlark rift (Fig. 1), from the relatively unextended Papuan peninsula in the south, across the D'Entrecasteaux MCCs, to the northern edge of the Trobriand platform. Data recorded by these stations provide, to our knowledge, the first available constraints on the deep structure of this rift.

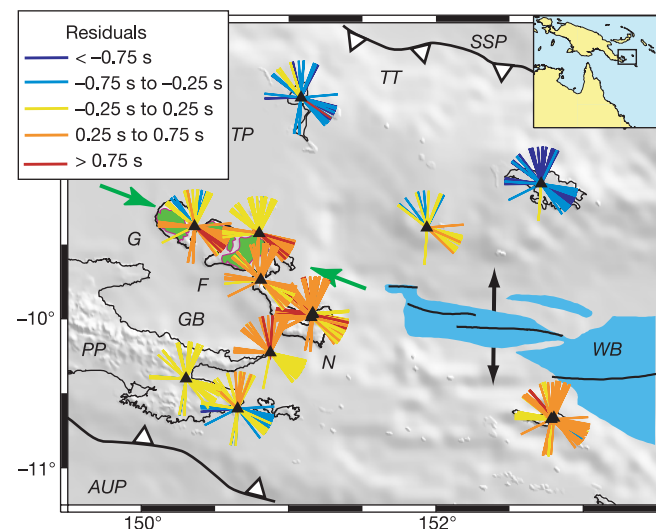


Figure 1 Tectonic features of the western Woodlark rift, showing seismic stations (small black triangles) and azimuths of incident teleseismic rays. Blue shaded region, new sea floor (< 2 Myr old); black arrows, direction of modern extension; green arrows, chain of MCCs. Letters denote the D'Entrecasteaux islands: Goodenough (G), Fergusson (F) and Normanby (N); the Papuan peninsula (PP), Goodenough basin (GB), Trobriand platform (TP), Trobriand trough (TT), Woodlark basin (WB), Australian plate (AUP), and Solomon Sea plate (SSP). Lines radiating from stations point at back-azimuth of incident teleseismic P waves, and indicate travel time residuals by colour. Blue lines, indicating negative residuals, show seismically fast regions while red regions indicate slow regions.

Teleseismic P coda constrains crustal thickness via receiver function techniques¹⁷. We generate receiver functions from the Woodlark broadband signals and invert for crustal structure (Figs 2 and 3; see Methods for details). The results document large variation in crustal thickness: beneath the relatively unextended Papuan peninsula, crustal thickness is 10–15 km greater than beneath the D'Entrecasteaux islands, 50–100 km to the north. The thinnest crust, 20 km thick, lies beneath Normanby island close to the oceanic rift tip, and crust remains thin beneath all D'Entrecasteaux stations. Crust thickens to the north, although a large gap in coverage obscures details of this transition, and the >40 km thickness at the northernmost station may include 6–8 km of oceanic crust underthrust southward from the Trobriand trough. Beneath some stations Moho depths show azimuthal variations in excess of 5 km, perhaps due to dipping Moho structure or unmodelled anisotropy; these complexities do not affect the overall Moho geometry, and so are not considered further. The regional pattern suggests that extension occurs where crust thins. This contrasts with the flat Moho beneath the transition between the Basin and Range and the Colorado Plateau³. Beneath the D'Entrecasteaux islands, the average Moho upwarp is 10–15 km, close to the amount of upper crust tectonically removed over the MCCs in the past 2 Myr, a correlation that suggests a causal link. Mid-crustal discontinuities could not be identified beneath most stations, so variations in lower-crustal thickness could not be directly measured.

The observed crustal thinning of 10–15 km should produce 1.5–3 km of subsidence, if isostatically balanced, depending upon whether or not the mantle lithosphere thins in concert with the crust (Fig. 4b). However, average elevations throughout the D'Entrecasteaux islands lie close to sea level, as do elevations at the end of the Papuan peninsula. Also, Bouguer gravity anomalies (Fig. 4c) indicate isostatic compensation. Hence, some other source of buoyancy must exist to support topography. To test the possibility that isostatic support comes from the mantle, we measure travel-time residuals in teleseismic P waves (Fig. 1). Residuals at stations on the north flanks of the MCCs show up to 1 s variation with azimuth, which requires a substantial velocity variation across this tectonic boundary (>5% if distributed uniformly over the upper 150 km); other stations on the MCCs are consistently slow,

requiring lower wave speeds in the mantle beneath the D'Entrecasteaux Islands than elsewhere. To illuminate the source of these residuals, we formally invert them for P velocities.

Inversion delineates slow velocities in the uppermost mantle associated with the D'Entrecasteaux islands and the oceanic rift, and faster velocities elsewhere (Figs 3 and 4a). These anomalies are strongest within the upper 100 km, with the uppermost mantle beneath the D'Entrecasteaux islands ~5% slower than surrounding material. Such variations are apparent in the pattern of residuals (Fig. 1), and show that at mantle depths the low-velocity region at the oceanic rift tip is contiguous with low velocities beneath these islands. The anomalies, if interpreted in terms of temperature alone, require 300–700 °C higher temperatures depending upon the effects of attenuation; the 700 °C value ignores possible physical dispersion effects¹⁸, and so provides an upper bound. Such high temperatures probably require melting. This is indeed likely, given the geologic evidence of Plio-Quaternary volcanism and granodiorite intrusion in the area¹⁹.

To describe the isostatic consequence of these anomalies, we compare the predicted gravity anomalies from both the crustal root and the mantle velocity structure to observed gravity (Fig. 4c). The only parameters adjusted to fit observed gravity are a constant term to account for long-wavelength effects, and two parameters describing plate flexure between the northernmost constraint and the Trobriand trough ($X = 170$ –250 km, where X is horizontal distance on Fig. 4). The predicted positive anomaly from the elevated Moho (replacing crust with mantle) shows the same amplitude and shape, but opposite sign, as that predicted from the mantle velocity structure. Hence, by Gauss's theorem, the excess mass associated with thinned crust across the D'Entrecasteaux islands equals the mass deficit of the upper mantle. No additional lateral density anomalies need be present to compensate topography. The observed gravity deviates by ~40 mGal from the predicted combination of both effects at wavelengths of ~250 km, perhaps reflecting dynamic effects, long-wavelength density variations, or errors in assumed velocity–density relationships (see Methods).

These observations show rifting at the onset of continental

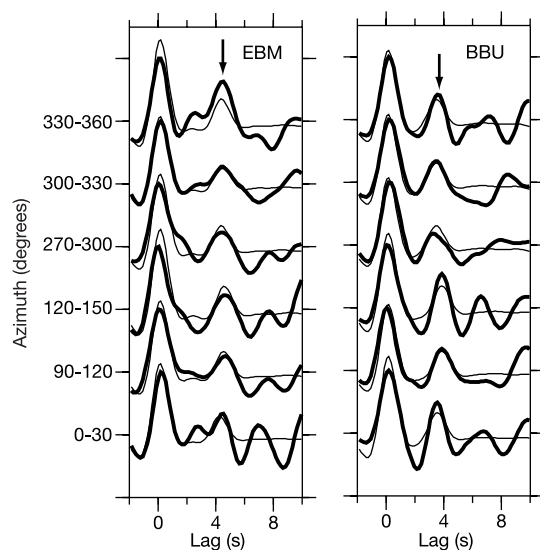


Figure 2 Receiver functions from two stations (EBM and BBU; see Fig. 3), stacked over 30° azimuthal bins. Thick lines, observed receiver functions; thin lines, best fit from inversion. The strong positive peak at 0 s corresponds to the direct P arrival, while the arrival at 3–5 s is the P-to-S conversion (Ps) off the Moho (arrow). Variations in Ps arrival time between stations reflect variation in crustal thickness.

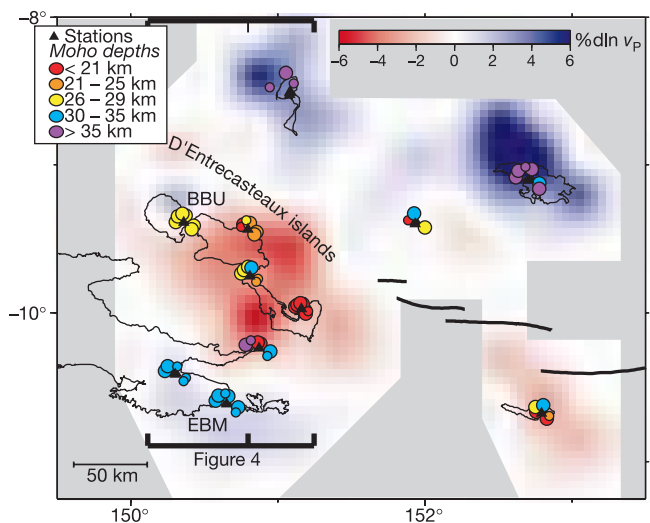


Figure 3 Imaging results beneath the western Woodlark rift. Colours show the percentage change to v_p from inversion of teleseismic P waves, at a depth of 55 km (uppermost mantle). Grey area, region of negligible resolution (resolution matrix diagonals <0.1). Coloured circles, depths to Moho estimated from receiver function inversions, plotted for each stack at average location of Ps conversion off Moho; smaller size indicates 95% confidence limits >10 km. Brackets, location of profile on Fig. 4. Triangles, seismic stations; thick line, location of active spreading segment. Note good correlation between slow regions (red) and thin crust, in particular the low-velocity region beneath the D'Entrecasteaux islands.

breakup in which MCCs form above the area of crustal thinning. In this region, mantle buoyancy and not crustal thickening isostatically supports regionally positive elevations for the D'Entrecasteaux islands region. Some second-order variations in lower-crustal thickness must explain along-strike variations in Moho depth beneath the MCCs, either through lower-crustal flow, magmatic addition, or variations in pre-extensional crustal thickness. Also, some flow of lower crust may be needed to explain subsidence and heat flow of basins adjacent to the D'Entrecasteaux islands²⁰. These along-strike changes in crustal thickness appear to be compensated through mantle heating, like the across-strike variations, as crustal thickness correlates with mantle velocities (Figs 1 and 3). Overall, the wholesale transfer of lower crust from unextended to highly

extended regions, inferred as the primary response to extension in the western US^{5,6}, need not be invoked here.

Several differences exist between the Woodlark rift and other rifts that may explain the juxtaposed thinning of crust and mantle lithosphere. Active continental rifts, such as the east Africa rift, similarly show large mantle heating^{13,21} but less than 20 km total surface extension, so provide little information on how the lithosphere responds to large extension. Only the western US and the Aegean region show comparable modern extension rates⁹, but both have relatively flat Mohos. One possibility is that pre-existing structure favours localization of crustal thinning at Woodlark, such as crust and mantle weakening from the Trobriand volcanic arc¹⁹. This feature would provide a zone of hot, weak sub-arc lithosphere to concentrate crustal thinning, MCC formation, and ultimately continental breakup. Additionally, time may play a factor. The relatively short time since the onset of exhumation at Woodlark may be insufficient for crustal flow, which probably occurs at timescales of the order of 2–10 Myr depending upon viscosities and channel thickness^{4–6}. This would explain why crust near the (older) MCCs has equilibrated, whereas crust in the young and active Woodlark rift has not. The Woodlark rift shows MCCs in development, before crustal flow has had a chance to complete.

The mantle velocities seen here suggest temperatures in excess of those predicted from pure-shear models of lithospheric thinning (Fig. 4b, orange line). (Hydration or enrichment from previous subduction may contribute to the low velocities here, although elevated temperatures and melt can explain most observations.) The P-wave anomalies locally exceed those found beneath mid-ocean ridges²², implying complete replacement of sub-continental mantle lithosphere by asthenosphere. Perhaps the pure-shear calculations underestimate mantle uplift because the pre-extensional crustal thickness beneath the MCCs may have been greater than in surrounding regions¹⁰, consistent with high pressures (5–6 kbar) at 4 Myr ago within the MCCs¹⁵. The additional mantle decompression may then suffice to produce buoyant melt as well as increased temperatures²³, further contributing to buoyancy. Such mantle heating is also a component of previous magmatism-driven models for MCC formation here²⁴, although such models elevate MCCs through crustal thickening, the opposite of what we observe.

As the mantle decompresses and approaches the peridotite solidus it should weaken, making this region susceptible to flow and further convective erosion of the mantle lithosphere beneath the MCCs²⁵. Thus, mantle lithosphere appears to have been largely removed beneath much of the D'Entrecasteaux islands over a short timescale, providing a 'snapshot' of processes likely to lead to the opening of an ocean basin. □

Methods

Receiver functions

Teleseismic P wavetrains impinging upon the crust below stations give rise to a sequence of P-to-S conversions (Ps) and reverberations, which are isolated via receiver function techniques to constrain structure¹⁷. We generate receiver functions using standard frequency-domain methods. We select up to 91 low-noise events per station at a wide range of azimuths, and stack the filtered receiver functions into azimuthal bins 30° wide (see Supplementary Information). At all stations we observe a prominent positive pulse at 3–6 s lag following the direct P at zero lag (Fig. 2), interpreted to be the Ps conversion off the Moho. Azimuthal stacks of these waveforms are inverted for one or two parameters, typically the Moho depth and mid-crustal interface depth, using a waveform fitting approach²⁶ to produce crustal thickness estimates with 95% confidence limits from *F* tests. These inversions give estimates of crustal thickness at several Ps conversion points at the Moho around each station. Conversion times are transformed into depths using a one-dimensional velocity structure determined by a joint inversion of arrival times and hypocentres from 30 well-recorded local earthquakes. The inversion gives crustal P velocities of 6.1 km s⁻¹ at depths >10 km, and upper-mantle P_n velocities of 7.8 km s⁻¹.

Travel time inversion

Residuals of teleseismic P waves are determined by waveform cross-correlation of the first cycle of incident P waveforms at frequencies of 0.15–1 Hz. From 102 events, 971 arrivals showed sufficient signal levels to use in inversion. Using standard methods²⁷ the differential residuals are inverted for velocity. Before inversion, the residuals are corrected

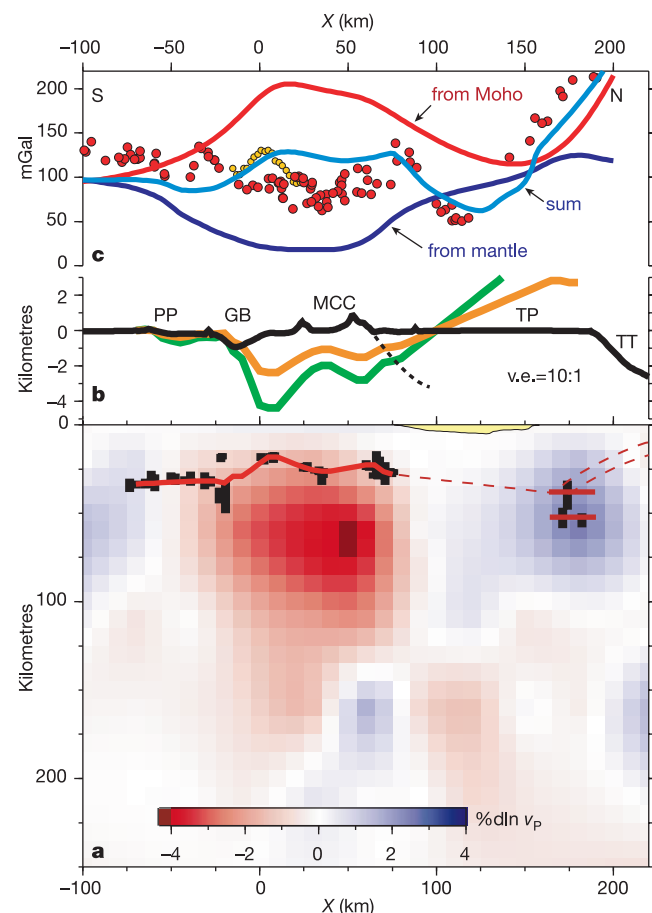


Figure 4 North-south cross-section through the D'Entrecasteaux MCCs at 150.8° E; horizontal distance $X = 0$ km corresponds to 10° S. **a**, Results of inversions. Colours represent perturbation to P velocities. Red line, Moho determined from receiver functions from stations west of 151.3° E smoothed at 10 km wavelength (and dashed where poorly constrained); vertical black bars, 95% confidence limits in individual measurements. For $X > 170$ km, inferred Moho depth is based on fit to gravity of a thin elastic plate representing Trobriand trough underthrusting. Yellow patch, Trobriand basin incorporated in final gravity model. **b**, Topography, observed (black line) and predicted (green, orange). Abbreviations same as Fig. 1. Green line, topography predicted from Moho relief alone, assuming constant crustal and mantle density. Orange line, predicted initial subsidence³⁰ assuming mantle lithosphere thinning beneath each point exactly matching the crustal thinning. **c**, Bouguer gravity anomalies. Red and yellow circles, values observed on land and at sea, respectively. Red line, anomaly predicted from crustal thickness changes; dark purple line, anomaly predicted from mantle velocity anomalies; light blue line, sum of both, plus effect of Trobriand basin. Note general agreement in region where Moho depths and mantle velocities are well constrained by seismic experiment ($-80 \text{ km} < X < 100 \text{ km}$). v.e., vertical exaggeration.

for variations in crustal thickness from receiver functions; velocities do not change in overall trend when corrections are not applied. The inversion is parameterized by a three-dimensional mesh spaced 30 km by 30 km horizontally and 50 km vertically, extending to 230 km depth. Experiments showed that variance reduction requires lateral heterogeneity only to 200 km depth, and that lateral resolution beneath the centre of the array is 30–40 km. In other words, the data set exhibits sensitivity to structures larger than 30–40 km within the upper 200 km of the Earth. Velocity anomalies in cross-section (Fig. 4) represent averages over three adjacent cells near 150.8° E longitude.

Gravity calculations

Gravity from variations in crustal thickness is calculated from the two-dimensional Moho geometry of Fig. 4a, assuming 400 kg m^{-3} Moho density contrast. North of the northernmost station ($X = 170 \text{ km}$), the Moho was forced to smoothly shallow to give 7-km-thick (oceanic) crust beneath the Trobriand trough ($X = 244 \text{ km}$) in a mathematical form consistent with elastic flexure. The gravitational effect of a low-density contrast (-300 kg m^{-3}) Trobriand basin was included ($X = 70\text{--}180 \text{ km}$), based on its known geometry²⁸. The anomaly due to mantle density variations ($\delta\rho$) was derived from the estimated three-dimensional velocity perturbations (δv_p), discretized into prisms of constant density. The conversion assumes $\delta \ln \rho / \delta \ln v_p = 0.54$, appropriate for dry pyrolyte at 2 GPa, based on a petrologically consistent database of mineral properties and equation-of-state calculations²⁹. This conversion does not include effects of physical dispersion from attenuation¹⁸.

Received 10 January; accepted 17 July 2002; doi:10.1038/nature00990.

1. Lister, G. S. & Davis, G. A. The origin of metamorphic core complexes and detachment faults formed during Tertiary continental extension in the northern Colorado River region, U.S.A. *J. Struct. Geol.* **11**, 65–94 (1989).
2. Klemperer, S. L., Hauge, T. A., Hauser, E. C., Oliver, J. E. & Potter, C. J. The Moho in the northern Basin and Range Province, Nevada, along the COCORP 40 degrees N seismic-reflection transect. *Geol. Soc. Am. Bull.* **97**, 603–618 (1986).
3. McCarthy, J., Larkin, S. P., Fuis, G. S., Simpson, R. W. & Howard, K. A. Anatomy of a metamorphic core complex; seismic refraction/wide-angle reflection profiling in southeastern California and western Arizona. *J. Geophys. Res.* **96**, 12259–12291 (1991).
4. Block, L. & Royden, L. H. Core complex geometries and regional scale flow in the lower crust. *Tectonics* **9**, 557–567 (1990).
5. Kruse, S., McNutt, M., Phipps-Morgan, J., Royden, L. & Wernicke, B. Lithospheric extension near Lake Mead: A model for ductile flow in the lower crust. *J. Geophys. Res.* **96**, 4435–4456 (1991).
6. McKenzie, D., Nimmo, E., Jackson, J. A., Gans, P. B. & Miller, E. L. Characteristics and consequences of flow in the lower crust. *J. Geophys. Res.* **105**, 11029–11046 (2000).
7. Lister, G. S. & Baldwin, S. L. Plutonism and the origin of metamorphic core complexes. *Geology* **21**, 607–610 (1993).
8. Davies, H. L. & Warren, R. G. Origin of eclogite-bearing, domed, layered metamorphic complexes ("core complexes") in the D'Entrecasteaux Islands, Papua New Guinea. *Tectonics* **7**, 1–21 (1988).
9. Abers, G. A. in *Non-Volcanic Rifting of Continental Margins: Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 305–318 (Geological Society, London, 2001).
10. Taylor, B., Goodliffe, A. M. & Martinez, F. How continents break up: Insights from Papua New Guinea. *J. Geophys. Res.* **104**, 7497–7512 (1999).
11. Weissel, J. K., Taylor, B. & Karner, G. D. The opening of the Woodlark Basin, subduction of the Woodlark spreading system, and the evolution of northern Melanesia since mid-Pliocene time. *Tectonophysics* **87**, 253–277 (1982).
12. Ruppel, C. Extensional processes in continental lithosphere. *J. Geophys. Res.* **100**, 24187–24215 (1995).
13. Nyblade, A. A., Owens, T. J., Gurrrola, H., Ritsema, J. & Langston, C. A. Seismic evidence for a deep mantle thermal anomaly beneath east Africa. *Geology* **28**, 599–602 (2000).
14. Baldwin, S. L., Lister, G. S., Hill, E. J., Foster, D. A. & McDougall, I. Thermochronologic constraints on the tectonic evolution of active metamorphic core complexes, D'Entrecasteaux Islands, Papua New Guinea. *Tectonics* **12**, 611–628 (1993).
15. Hill, E. J. & Baldwin, S. L. Exhumation of high pressure metamorphic rocks during crustal extension in the D'Entrecasteaux Islands, Papua New Guinea. *J. Metamorph. Geol.* **11**, 261–277 (1993).
16. Ollier, C. D. & Pain, C. F. Active rising surficial gneiss domes in Papua New Guinea. *J. Geol. Soc. Aust.* **27**, 33–44 (1980).
17. Owens, T. J., Zandt, G. & Taylor, S. R. Seismic evidence for an ancient rift beneath the Cumberland Plateau, TN: a detailed analysis of broadband teleseismic P-waveforms. *J. Geophys. Res.* **89**, 7783–7795 (1984).
18. Karato, S. Importance of anelasticity in the interpretation of seismic tomography. *Geophys. Res. Lett.* **20**, 1623–1626 (1993).
19. Smith, I. E. in *Volcanism in Australasia* (ed. Johnson, R. W.) 275–286 (Elsevier, New York, 1976).
20. Martinez, F., Goodliffe, A. M. & Taylor, B. Metamorphic core complex formation by density inversion and lower-crust extrusion. *Nature* **411**, 930–934 (2001).
21. Achauer, U. et al. New ideas on the Kenya rift based on the inversion of the combined dataset of the 1985 and 1989/1990 seismic tomography experiments. *Tectonophysics* **238**, 305–329 (1994).
22. Toomey, D. R., Wilcock, W. S., Solomon, S. C., Hammond, W. C. & Orcutt, J. A. Mantle seismic structure beneath the MELT region of the East Pacific Rise from P and S wave tomography. *Science* **280**, 1224–1227 (1998).
23. McKenzie, D. & Bickle, M. J. The volume and composition of melt generated by extension of lithosphere. *J. Petrol.* **29**, 625–679 (1988).
24. Hill, E. J., Baldwin, S. L. & Lister, G. S. Magmatism as an essential driving force for formation of active metamorphic core complexes in eastern Papua New Guinea. *J. Geophys. Res.* **100**, 10441–10451 (1995).
25. Buck, W. R. & Su, W. Focused mantle upwelling below mid-ocean ridges due to feedback between viscosity and melting. *Geophys. Res. Lett.* **16**, 641–644 (1989).
26. Sheehan, A. F., Abers, G. A., Jones, C. H. & Lerner-Lam, A. L. Crustal thickness variations across the Colorado Rocky Mountains from teleseismic receiver functions. *J. Geophys. Res.* **100**, 20391–20404 (1995).

27. Roecker, S. W. in *Seismic Tomography Theory and Practice* (eds Iyer, H. M. & Hirahara, K.) 534–611 (Chapman and Hall, London, 1993).
28. Pinchin, J. & Bembrick, C. Cape Vogel Basin, PNG—tectonics and petroleum potential. *Bur. Min. Resour. Aust. Rec.* **32**, 31–37 (1985).
29. Hacker, B. R., Abers, G. A. & Peacock, S. M. Subduction factory 1: Theoretical mineralogy, density, seismic wavespeeds, and H₂O content. *J. Geophys. Res.* (in the press).
30. McKenzie, D. P. Some remarks on the development of sedimentary basins. *Earth Planet. Sci. Lett.* **40**, 25–32 (1978).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

The field effort in PNG would not have been successful without the efforts of S. Sioni, University of Papua New Guinea, J. Dickson, H. Sailasa and others of the Milne Bay Provincial Administration, PNG, and sub-district administrators throughout the province. Instruments and technical assistance were made available through the IRIS PASSCAL Instrument Center. The Geological Survey of Papua New Guinea gave permission to publish the gravity data on Fig. 4. We thank C. Ruppel for comments. This work was supported by the Geosciences Directorate of the National Science Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.A. (e-mail: abers@bu.edu).

A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*

Bryony A. P. Williams*, Robert P. Hirt*, John M. Lucocq† & T. Martin Embley*

* Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

† School of Life Sciences, WTB/MSI complex, University of Dundee, Dundee DD1 5EH, UK

Microsporidia are obligate intracellular parasites of several eukaryotes. They have a highly complex and unique infection apparatus but otherwise appear structurally simple¹. Microsporidia are thought to lack typical eukaryotic organelles, such as mitochondria and peroxisomes. This has been interpreted as support for the hypothesis that these peculiar eukaryotes diverged before the mitochondrial endosymbiosis, which would make them one of the earliest offshoots in eukaryotic evolution^{2,3}. But microsporidian nuclear genes that encode orthologues of typical mitochondrial heatshock Hsp70 proteins have been detected, which provides evidence for secondary loss of the organelle or endosymbiont^{4–6}. In addition, gene trees and more sophisticated phylogenetic analyses have recovered microsporidia as the relatives of fungi, rather than as basal eukaryotes^{7–9}. Here we show that a highly specific antibody raised against a *Trachipleistophora hominis* Hsp70 protein detects the presence, under light and electron microscopy, of numerous tiny (~50 × 90 nm) organelles with double membranes in this human microsporidian parasite. The finding of relictual mitochondria in microsporidia provides further evidence of the reluctance of eukaryotes to lose the mitochondrial organelle, even when its canonical function of aerobic respiration has been apparently lost.

Molecular markers attest to a genetic heritage of the α-proteobacterial mitochondrial endosymbiont in microsporidia^{4–6,10,11}; however, no experimental link has been shown between any of

these genetic markers, the production of a protein and its targeting to an organelle. Genes encoding the α - and β -subunits of pyruvate dehydrogenase have been found in *Nosema locustae* and *Encephalitozoon cuniculi*^{11,12}. In aerobes these proteins, which are also of endosymbiotic origin¹¹, are encoded in the nucleus and are imported into mitochondria using an amino-terminal cleaved transit peptide¹³. But the microsporidian genes do not encode a recognizable mitochondrial targeting signal and their cellular location is unknown¹¹.

Genes that encode homologues of mitochondrial Hsp70 (mtHsp70) have been found in the genomes of all microsporidia investigated so far^{4-6,10}. A mtHsp70 on the nuclear genome of eukaryotes has been proposed to result from horizontal gene transfer from the α -proteobacterial endosymbiont genome, early in the history of the mitochondrial endosymbiosis. In aerobic eukaryotes this protein is involved in various essential functions in the mitochondrion^{14,15}, including the translocation of nuclear-

encoded mitochondrial proteins and the folding of iron-sulphur (Fe-S) clusters^{13,16}. Thus, of all of the mitochondrial endosymbiont-derived genes detected in microsporidian genomes, mtHsp70 provides the strongest indication that an organelle might still be present.

To test this hypothesis, we cloned and sequenced a gene encoding a mtHsp70 from *T. hominis*. This microsporidian was isolated from an individual affected with AIDS and is perpetuated in a rabbit kidney cell line (RK13)¹⁷, which provides one of the few tractable systems for biochemical and cellular investigations in microsporidia. The *T. hominis* mtHsp70 contains the key residues that are required for ATPase activity (Fig. 1a), and the gene is expressed as messenger RNA and protein in spores and intracellular stages of the parasite (Fig. 2). Detailed phylogenetic analysis of the DNA (Fig. 1c) and protein sequence (not shown) showed that it clusters with other microsporidian mtHsp70s, and together these form a strongly supported monophyletic group with fungal mtHsp70. The topology of the mtHsp70 tree is consistent with published species relationships⁷⁻⁹, which suggests that the genes for microsporidian mtHsp70 were inherited by vertical descent.

To investigate the cellular expression of the *T. hominis* mtHsp70 protein *in situ*, we carried out immunolocalization experiments on parasite-infected rabbit kidney cells using a specific antibody against mtHsp70. This is a polyclonal rabbit antibody raised against a *T. hominis* mtHsp70 fusion protein produced in *Escherichia coli*. The mtHsp70 protein was only detected in microsporidia (Fig. 3), and there were no significant cross-reactions with host proteins, which was consistent with the results of western blots (Fig. 2).

Staining was limited to small discrete structures distributed throughout the cytoplasm of the wall-less meront form of the parasite. The thick walls of the sporophorous vesicles and spores were apparently impermeable to the antibody under the conditions used. Meronts of *T. hominis* contained between 7 and 47 structures with a mean $\pm$ s.d. of 28 ± 13 structures ($n = 11$). The size of the antibody-positive structures could not be determined reliably because they were close to the resolution limit (~ 200 nm) of confocal microscopy. The labelled structures were not stained by Mitotracker Red (data not shown), which detects the membrane potential generated by the typical mitochondria of the host. The genome sequence of *E. cuniculi*¹² lacks genes for the mitochondrial electron transport chain and the F_0F_1 ATPase complex, therefore, this lack of detectable membrane potential in the microsporidian organelles is not surprising.

We used electron microscopy to determine whether the mtHsp70 was localized to a membrane-bound compartment. This was particularly pertinent because we could not identify an N-terminal

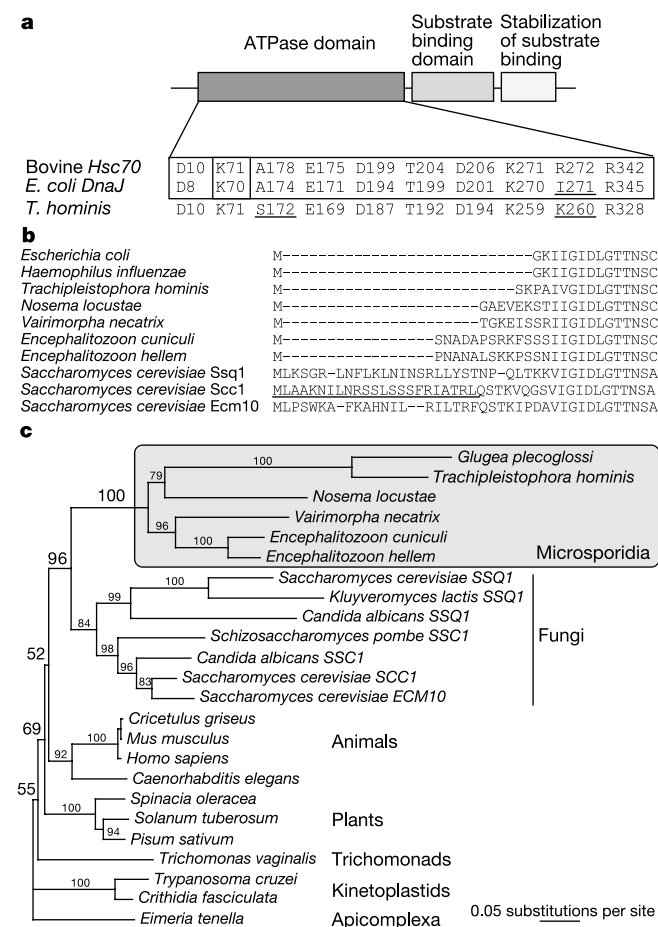


Figure 1 Characterization of a *T. hominis* gene encoding a mitochondrial Hsp70. **a**, Structural and functional organization of Hsp70 proteins. Residues for which mutagenesis affects the ATPase activity in *E. coli* DnaJ and/or bovine Hsc70 proteins are shown in boxes above the corresponding *T. hominis* sequence. Mutagenesis of K71/70 eliminates ATPase activity^{28,29}. Variable residues are underlined. **b**, N-terminal sequences of microsporidian Hsp70 showing the absence of a recognizable mitochondrial transit peptide; the three yeast mtHsp70 paralogues are shown for comparison. The experimentally determined cleaved transit peptide for the yeast Ssc1 protein is underlined. **c**, Phylogenetic analysis of the *T. hominis* Hsp70 DNA sequence showing its relationship to other microsporidian Hsp70 gene sequences and to mitochondrial orthologues. The tree shown is a consensus tree calculated from 1,000 bootstrap replicates using DNA variable codon positions 1 and 2 (1,338 sites), the LogDet/Paralinear distance method, and minimum evolution. Bootstrap values are given at relevant nodes.

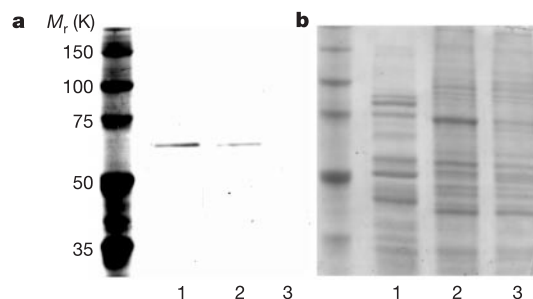


Figure 2 Analysis of the expression of the *T. hominis* Hsp70 in spores and infected and non-infected rabbit kidney (RK13) cells using a rabbit polyclonal antibody against *T. hominis* Hsp70. **a**, Western blots of total protein extracts from *T. hominis* spores (lane 1), *T. hominis* infected RK cells (lane 2) and non-infected RK13 cells (lane 3). A protein with an apparent relative molecular mass (M_r) of about 60,000 (60K) was detected in spores and in *T. hominis* infected RK13 cells but not in uninfected RK cells. The size of this band corresponds to the theoretical M_r deduced from the gene sequence (60K). **b**, Coomassie staining of a duplicate gel showing total proteins.

extension in the *T. hominis* mtHsp70 that might target it to an organelle (Fig. 1b), nor did we detect any targeting to mammalian mitochondria in heterologous transfection experiments (data not shown). Immunogold labelling of *T. hominis* meronts (and sporonts) in rabbit kidney cells showed that mtHsp70 labelling was concentrated inside elongated uniformly sized structures, which in sections measured roughly 50 nm by 90 nm (Fig. 4). These structures were surrounded by a double membrane, and most gold particles were located over the matrix at the inner aspect of the inner

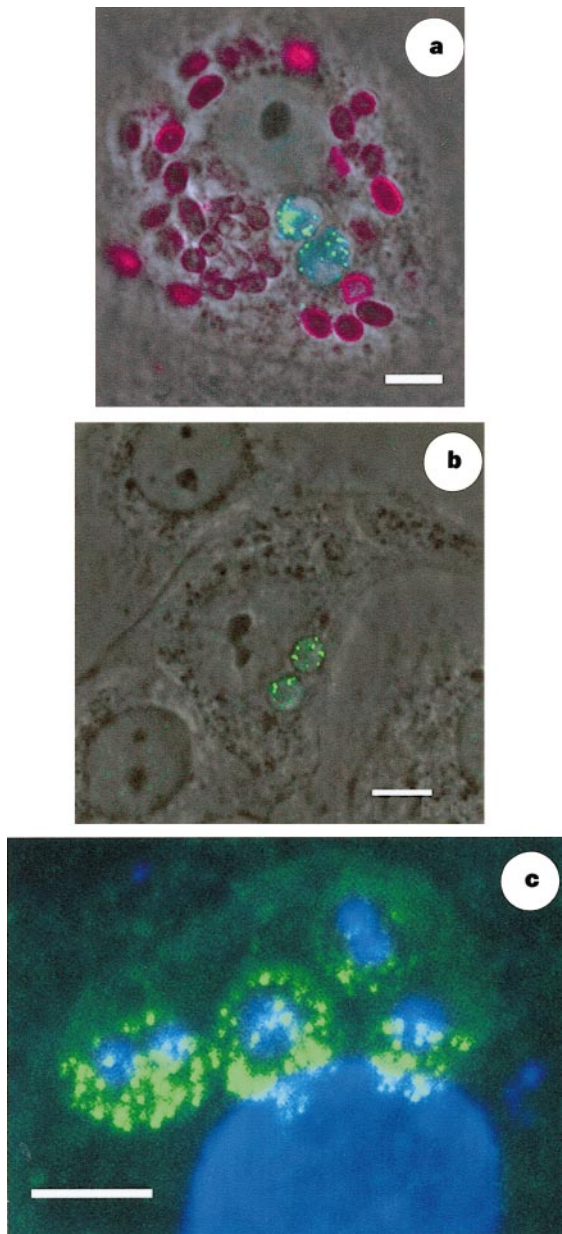


Figure 3 Cellular localization of the *T. hominis* Hsp70 protein detected *in situ* using immunofluorescence and confocal microscopy. **a**, A heavily infected rabbit kidney cell containing several spores (stained red) and two wall-less meronts of *T. hominis*. Spores of *T. hominis* were detected using a mouse polyclonal antibody raised against spore proteins (red). The meronts contain numerous small structures (green) that react specifically with the polyclonal antibody against *T. hominis* mtHsp70. **b**, Two meronts containing mtHsp70-positive structures in a field of otherwise uninfected rabbit kidney cells. **c**, Composite reconstructed image of four meronts, which contain several mtHsp70-positive structures, inside a single infected cell. The nuclei of host and parasite were labelled with DAPI (blue). Scale bar, 5 μ m.

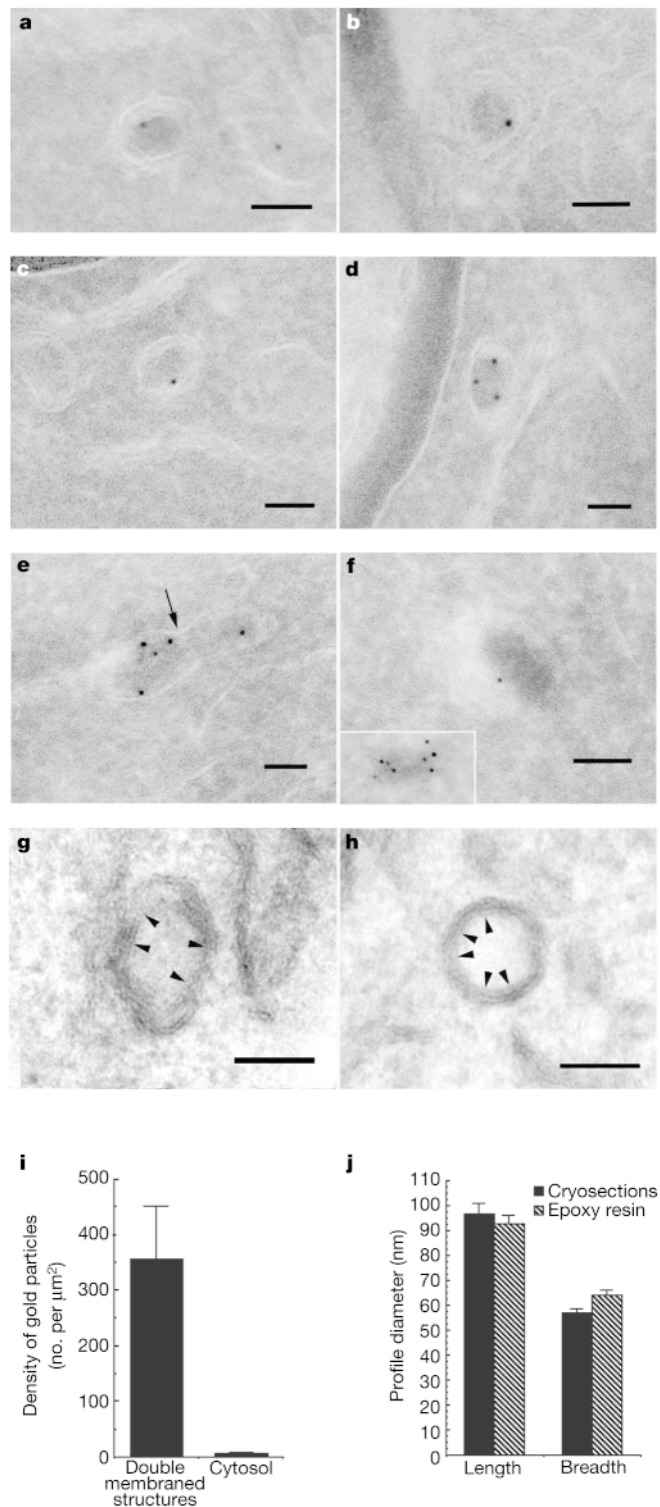


Figure 4 Transmission electron microscopy of microsporidian structures. **a–e**, Thawed frozen sections of glutaraldehyde-fixed parasites showing evidence of the double membrane and association of gold particles with the inner aspect of the inner membrane (**a–d**) and a putative dividing form (**e**; septa indicated by an arrow). **f**, Characteristic electron density of the inner matrix after fixation with formaldehyde/glutaraldehyde mixtures (membranes are less evident with this fixation); inset shows an intensely labelled profile obtained for formaldehyde alone. Shown are meronts (**a**, **c–f**) and a sporont (**b**). **g**, **h**, Double membranes are clearly identified for structures in meronts after embedding in epoxy resin (inner membrane indicated by arrowheads). **i**, Quantification of the gold labelling density (mean of two experiments; error bars indicate range). **j**, Sizes of double-membrane structures in thawed frozen sections and epoxy resin sections (cryosections, $n = 21$; epoxy resin, $n = 18$). Error bars indicate s.e.m. Scale bar, 50 nm.

membrane (66% of total, $n = 50$ gold particles). This is consistent with the suggested functional location of mtHsp70 in the mitochondria of other eukaryotes¹³. The matrix was characteristically denser after fixation in formaldehyde/glutaraldehyde mixtures (Fig. 4f). After formaldehyde fixation alone, the matrix was intensely labelled with up to eight particles per structure (Fig. 4f, inset), although the membranes were not as easily seen. Structures with double membranes of similar size were observed after conventional electron microscopy processing (Fig. 4g, h). The approximate number of double-membrane organelles was estimated as 19 per cell—a value that is consistent with the counts obtained by confocal microscopy.

The function of the new organelle and its importance to microsporidia is unknown, but the genome of *E. cuniculi* lacks the genes for aerobic respiration¹². The organelles are relatively abundant, which is consistent with an active role in the metabolism of the parasite. The genome of *E. cuniculi* contains five genes of α -proteobacterial ancestry that, in other eukaryotes, function in mitochondrial Fe-S cluster assembly^{12,18}. The generation of Fe-S clusters, in which mtHsp70 is involved, is currently the only biosynthetic mitochondrial process known to be essential to the yeast cell¹⁶. On the basis of the presence of these five genes, the existence in microsporidia of an undiscovered organelle of mitochondrial ancestry involved in Fe-S cluster assembly was previously predicted¹². We agree that this is a plausible selection pressure for retaining the microsporidian organelle, for which we have now provided physical evidence; however, we could not confidently detect potential mitochondrial-like targeting motifs encoded by these genes using *in silico* analyses, and there are no experimental data for the location of their proteins in microsporidia. As our results indicate, even in the absence of a detectable transit peptide, the *T. hominis* mtHsp70 is localized to a double-membrane-bound organelle. Microsporidia are well known to have streamlined much of their biology in adapting to their parasitic lifestyle^{1,12}. It is possible that this process has extended to the parasite mitochondrial protein import machinery.

It is apparent that many microbial eukaryotes that were previously thought to lack mitochondria have actually retained the organelle. There is good evidence that hydrogenosomes of phylogenetically diverse eukaryotes, including anaerobic ciliates¹⁹, fungi²⁰ and trichomonads²¹, are mitochondria that have lost the capacity for aerobic respiration and gained the biochemistry to make hydrogen. *Entamoeba histolytica*, an anaerobic amoeba, contains an organelle called either a mitosome to denote its mitochondrial ancestry²², or a crypton²³ to denote its still unknown functions. We have shown the obligately parasitic microsporidia also contain a mitochondrial remnant. Its diminutive size and the absence of a reliable marker have almost certainly contributed to a failure to detect it previously, despite microsporidia being studied extensively by electron microscopy^{1,24}. Our findings suggest that other eukaryotes that are currently thought to lack mitochondria such as *Giardia*, which contains mitochondrial chaperonin 60 (ref. 25), may be shown to contain an organelle of mitochondrial ancestry. □

Methods

Growth of *T. hominis* in rabbit kidney cells

Rabbit kidney cells were infected and grown as described¹⁷.

Amplification of *T. hominis* mtHsp70

We designed degenerate Hsp70 primers from an alignment of mitochondrial Hsp70 proteins⁵ and used them to amplify 1,016 base pairs from DNA extracted from purified spores. The 5' and 3' ends of the gene were amplified using the Marathon Complementary DNA Amplification kit (Clontech) using *T. hominis* spore mRNA or infected cell-culture mRNA as a template. Amplified fragments were sequenced using the BigDye sequencing kit (Amersham). We confirmed the sequence by amplification and sequencing of the gene from genomic DNA.

Recombinant protein and antibody preparation

We cloned the full-length *T. hominis* mtHsp70 gene into the pET30a expression vector

(Novagen). BL2 DE3 Lysogen *E. coli* cells (Novagen) were transformed with the vector and protein expression was induced. The protein was extracted from inclusion bodies using the Bugbuster protein extraction reagent (Novagen), purified by gel electrophoresis and extracted from the gel by electroelution. We sent 1.2 mg of purified protein to Harlan Sera Labs for the production of a rabbit polyclonal antibody.

Western blotting

Whole-cell extracts were prepared from *T. hominis* spores, rabbit kidney cells infected with *T. hominis* and uninfected rabbit kidney cells. Extracts were blotted and blocked using standard protocols. Blots were incubated with the antibody at a dilution of 1:1,000 and then with a secondary anti-rabbit antibody conjugated to alkaline phosphatase (Sigma). We developed the blots with Sigmafast (Sigma).

Immunolocalization of mtHsp70 in cells infected by *T. hominis*

Microsporidia-infected rabbit kidney cells were grown on cover slips until confluence and then fixed in 50:50 acetone:methanol (v/v) at -20°C for 2 h. Blocked cells were incubated in PBS, 1% milk and a 1:200 dilution of the rabbit antibody against *T. hominis* mitochondrial Hsp70 for 1 h. In some experiments, a 1:200 dilution of a mouse antibody against *T. hominis* spores was also added²⁶. We incubated the slides with goat secondary antibodies conjugated to fluorescent dye: the anti-mouse antibody was labelled with Alexa-fluor 568 and the anti-rabbit antibody with Alexa-fluor 488 (Molecular Probes). Labelled cells were visualized using a Leica SP confocal fluorescent microscope or a Zeiss 510 confocal laser scanning microscope for 4',6'-diamidino-2-phenylindole dihydrochloride (DAPI)-stained cells. To facilitate the counting of mtHsp70-positive structures, we photographed meronts in 20 confocal sections, which were then stacked on top of each other.

Immunoelectron microscopy

Parasite-infected monolayer RK13 cells were fixed in 0.5% glutaraldehyde or 4% paraformaldehyde/0.1% glutaraldehyde in 0.2 M PIPES (pH 7.2) for 30 min at room temperature, scraped, pelleted and cryoprotected in 2.1 M sucrose in PBS, frozen on cryosectioning stubs in liquid nitrogen, and cryosectioned at 50–100 nm on a Leica Ultracut UCT fitted with an EMFCS cryochamber. Some samples were fixed in 4% paraformaldehyde in 0.2 M PIPES (pH 7.2) but left in fixative until cryoprotection. Thawed cryosections were labelled using rabbit antisera or purified immunoglobulin- γ against mtHsp70, followed by 8-nm particles of protein-A-gold, contrasted in methylcellulose/uranyl acetate, and observed in a JEOL 1200EX. For quantification, elongated or round profiles with evidence of double membranes and measuring $<0.5\ \mu\text{m}$ (largest diameter) were photographed at a nominal magnification of $\times 40,000$. Micrographs were scanned at 1,000 d.p.i. and further magnified $\times 7.9$ in Adobe Photoshop version 5.5. Square lattice grids with spacings of 25 nm and 200 nm were generated on-screen and point counting was used to estimate the areas of double-membrane profiles and cytosol, respectively (area = sum of point counts $\times$ area associated with one point). We measured calliper diameters (length, longest axis and breadth, orthogonal to longest axis) on double-membrane profiles, which, in the case of thawed cryosections, were also labelled with immunogold.

Sequence alignment and phylogenetic analysis

We used CLUSTALW to align the *T. hominis* Hsp70 amino acid sequence to 23 other mitochondrial Hsp70 sequences. Further refinements were made by eye in the program GDE. A DNA alignment was created from the amino acid alignment. A LogDet and minimum evolution analysis was used to estimate the tree²⁷. Before analysis, 29.6% of sites were estimated by maximum likelihood as being unable to vary and were thus removed to correct for rate heterogeneity between sequences⁸.

Received 19 March; accepted 11 June 2002; doi:10.1038/nature00949.

- Vavra, J. & Larsson, J. I. R. In *The Microsporidia and Microsporidiosis* (eds Wittner, M. & Weiss, L. M.) 7–84 (American Society for Microbiology, Washington DC, 1999).
- Cavalier-Smith, T. Eukaryotes with no mitochondria. *Nature* **326**, 332–333 (1987).
- Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. & Woese, C. R. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature* **326**, 411–414 (1987).
- Peyretailade, E. *et al.* Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. *Mol. Biol. Evol.* **15**, 683–689 (1998).
- Hirt, R. P., Healy, B., Vossbrinck, C. R., Canning, E. U. & Embley, T. M. A mitochondrial Hsp70 orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria. *Curr. Biol.* **7**, 995–998 (1997).
- Germot, A., Philippe, H. & Le Guyader, H. Evidence for loss of mitochondria in Microsporidia from a mitochonondrial-type HSP70 in *Nosema locustae*. *Mol. Biochem. Parasitol.* **87**, 159–168 (1997).
- Keeling, P. J., Luker, M. A. & Palmer, J. D. Evidence from β -tubulin phylogeny that microsporidia evolved from within the fungi. *Mol. Biol. Evol.* **17**, 23–31 (2000).
- Hirt, R. P. *et al.* Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. *Proc. Natl Acad. Sci. USA* **96**, 580–585 (1999).
- Van de Peer, Y., Ben Ali, A. & Meyer, A. Microsporidia: accumulating molecular evidence that a group of amitochondriate and supposedly primitive eukaryotes are just curious fungi. *Gene* **246**, 1–8 (2000).
- Ariseu, N., Sanchez, L. B., Weiss, L. M., Müller, M. & Hashimoto, T. Mitochondrial-type hsp70 genes of the amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica* and two microsporidians. *Parasitol. Int.* **51**, 9–16 (2002).
- Fast, N. M. & Keeling, P. J. α and β subunits of pyruvate dehydrogenase E1 from the microsporidian *Nosema locustae*: mitochondrion-derived carbon metabolism in microsporidia. *Mol. Biochem. Parasitol.* **117**, 201–209 (2001).
- Katinka, M. D. *et al.* Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414**, 450–453 (2001).

13. Pfanner, N. & Geissler, A. Versatility of the mitochondrial protein import machinery. *Nature Rev. Mol. Cell. Biol.* **2**, 339–349 (2001).
14. Winzler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
15. Bukau, B. & Horwich, A. L. The Hsp70 and Hsp60 chaperone machines. *Cell* **92**, 351–366 (1998).
16. Lill, R. & Kispal, G. Maturation of cellular Fe-S proteins: an essential function of mitochondria. *Trends Biochem. Sci.* **25**, 352–356 (2000).
17. Hollister, W. S. *et al.* Development and ultrastructure of *Trachipleistophora hominis* n.g., n.sp. after *in vitro* isolation from an AIDS patient and inoculation into athymic mice. *Parasitology* **112**, 143–154 (1996).
18. Huynen, M. A., Snel, B., Bork, P. & Gibson, T. J. The phylogenetic distribution of frataxin indicates a role in iron-sulfur cluster protein assembly. *Hum. Mol. Genet.* **10**, 2463–2468 (2001).
19. Akhmanova, A. *et al.* A hydrogenosome with a genome. *Nature* **396**, 527–528 (1998).
20. van Der Giezen, M. *et al.* Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. *EMBO J.* **21**, 572–579 (2002).
21. Dyall, S. D. & Johnson, P. J. Origins of hydrogenosomes and mitochondria: evolution and organelle biogenesis. *Curr. Opin. Microbiol.* **3**, 404–411 (2000).
22. Tovar, J., Fischer, A. & Clark, C. G. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol. Microbiol.* **32**, 1013–1021 (1999).
23. Mai, Z. *et al.* Hsp60 is targeted to a cryptic mitochondrion-derived organelle ('crypton') in the microaerophilic protozoan parasite *Entamoeba histolytica*. *Mol. Cell. Biol.* **19**, 2198–2205 (1999).
24. Canning, E. U. In *Handbook of Protozoa* (eds Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J.) 53–72 (Jones and Bartlett, Boston, 1990).
25. Roger, A. J. *et al.* A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc. Natl Acad. Sci. USA* **95**, 229–234 (1998).
26. Cheney, S. A., Lafranchi-Tristem, N. J. & Canning, E. U. Serological differentiation of microsporidia with special reference to *Trachipleistophora hominis*. *Parasite* **8**, 91–97 (2001).
27. Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11**, 605–612 (1994).
28. Wilbanks, S. In *Guidebook to the Molecular Chaperones and Protein-Folding Catalysts* (ed. Gething, M.-J. H.) 13–18 (Oxford Univ. Press, New York, 1997).
29. Buchberger, A. & Bukau, B. In *Guidebook to the Molecular Chaperones and Protein-Folding Catalysts* (ed. Gething, M.-J. H.) 22–25 (Oxford Univ. Press, New York, 1997).

Acknowledgements

We thank E. Canning for the *T. hominis* culture system and advice on growing microsporidia; M. Duchon and A. Ball for help with confocal microscopy; C. Thomson for help with electron microscopy; and C. Danpure and G. Birdsey for comments on the manuscript and help with heterologous transfection experiments. J.M.L. was supported by a Wellcome Trust Research Leave Fellowship and Tenovus Scotland. B.A.P.W. was supported by a Wellcome Trust Biodiversity studentship, R.P.H. was supported by a Wellcome Trust University award.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.M.E. (e-mail: tme@nhm.ac.uk). The sequence of *T. hominis* mtHsp70 has been deposited in GenBank under accession code AF492453.

Molecular evolution of *FOXP2*, a gene involved in speech and language

Wolfgang Enard*, Molly Przeworski*, Simon E. Fisher†, Cecilia S. L. Lai†, Victor Wiebe*, Takashi Kitano*, Anthony P. Monaco† & Svante Pääbo*

* Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany

† Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

Language is a uniquely human trait likely to have been a prerequisite for the development of human culture. The ability to develop articulate speech relies on capabilities, such as fine control of the larynx and mouth¹, that are absent in chimpanzees and other great apes. *FOXP2* is the first gene relevant to the human ability to develop language². A point mutation in *FOXP2* co-segregates with a disorder in a family in which half of the members have severe articulation difficulties accompanied by linguistic and grammatical impairment³. This gene is disrupted by translocation in an unrelated individual who has a similar

disorder. Thus, two functional copies of *FOXP2* seem to be required for acquisition of normal spoken language. We sequenced the complementary DNAs that encode the *FOXP2* protein in the chimpanzee, gorilla, orang-utan, rhesus macaque and mouse, and compared them with the human cDNA. We also investigated intraspecific variation of the human *FOXP2* gene. Here we show that human *FOXP2* contains changes in amino-acid coding and a pattern of nucleotide polymorphism, which strongly suggest that this gene has been the target of selection during recent human evolution.

FOXP2 (forkhead box P2) is located on human chromosome 7q31, and its major splice form encodes a protein of 715 amino acids belonging to the forkhead class of transcription factors². It contains a glutamine-rich region consisting of two adjacent polyglutamine tracts, encoded by mixtures of CAG and CAA repeats. Such repeats are known to have elevated mutation rates. In the case of *FOXP2*, the lengths of the polyglutamine stretches differed for all taxa studied. Variation in the second polyglutamine tract has been observed in a small family affected with speech and language impairment, but this did not co-segregate with disorder, suggesting that minor changes in length may not significantly alter the function of the protein⁴. If the polyglutamine stretches are disregarded, the human *FOXP2* protein differs at only three amino-acid positions from its orthologue in the mouse (Fig. 1). When compared with a collection of 1,880 human–rodent gene pairs⁵, *FOXP2* is among the 5% most-conserved proteins. The chimpanzee, gorilla and rhesus macaque *FOXP2* proteins are all identical to each other and carry only one difference from the mouse and two differences from the human protein, whereas the orang-utan carries two differences from the mouse and three from humans (Fig. 1). Thus, although the *FOXP2* protein is highly conserved, two of the three amino-acid differences between humans and mice occurred on the human lineage after the separation from the common ancestor with the chimpanzee. These two amino-acid differences are both found in exon 7 of the *FOXP2* gene and are a threonine-to-asparagine and an asparagine-to-serine change at positions 303 and 325, respectively. Figure 2 shows the amino-acid changes, as well as the silent changes, mapped to a phylogeny of the relevant primates.

We compared the *FOXP2* protein structures predicted by a variety of methods⁶ for humans, chimpanzees, orang-utans and mice. Whereas the chimpanzee and mouse structures were essentially identical and the orang-utan showed only a minor change in secondary structure, the human-specific change at position 325 creates a potential target site for phosphorylation by protein kinase C together with a minor change in predicted secondary structure. Several studies have shown that phosphorylation of forkhead transcription factors can be an important mechanism mediating transcriptional regulation^{7,8}. Thus, although the *FOXP2* protein is extremely conserved among mammals, it acquired two amino-acid changes on the human lineage, at least one of which may have functional consequences. This is an intriguing finding, because

Table 1 Variation at the *FOXP2* locus in humans

No. of chromosomes sequenced	40
Length covered (double stranded, all individuals)	14,063 bp
Divergence from the chimp sequence*	0.87%
No. of variable positions	47
Singletons (no. of variable sites occurring at frequency 1 and 39)	31
θ_w (nucleotide diversity based on the no. of polymorphic sites)	0.079%
θ_π (mean nucleotide diversity)	0.03%
θ_H (nucleotide diversity with more weight given to alleles at high frequency ¹⁷)	0.117%
D ($P < 0.01$)†	–2.20
H ($P < 0.05$)‡	–12.24

*The corresponding value for the orang-utan is 2.5.

†A negative D value indicates a relative excess of low-frequency alleles¹⁶.

‡A negative H value indicates a relative excess of high-frequency derived alleles¹⁷.

FOXP2 is the first gene known to be involved in the development of speech and language.

To investigate whether the amino acids encoded in exon 7 are polymorphic in humans, we sequenced this exon from 44 human chromosomes originating from all major continents. In no case was any amino-acid polymorphism found. Further, a study that analysed the complete coding region of *FOXP2* in 91 unrelated individuals of mainly European descent found no amino-acid replacements except for one case of an insertion of two glutamine codons in the second polyglutamine stretch⁴. Because the two amino-acid variants specific to humans occur in 226 human chromosomes, this suggests that they are fixed among humans.

The evolutionary lineages leading to humans and mice diverged about 70 million years (Myr) ago^{9,10}. Thus, during the roughly 130 Myr of evolution that separate the common ancestor of humans and chimpanzees from the mouse, a single amino-acid change occurred in the *FOXP2* protein. By contrast, since the human and chimpanzee lineages diverged about 4.6–6.2 Myr ago¹¹, two fixed amino-acid changes occurred on the human lineage whereas none occurred on the chimpanzee and the other primate lineages, except for one change on the orang-utan lineage. We used a likelihood ratio¹² to test for constancy of the ratio of amino-acid replacements over nucleotide changes that do not cause amino-acid changes among the evolutionary lineages in Fig. 2. Whereas a significant increase in this ratio was observed on the human lineage

($P < 0.001$), no such increase was seen on any other lineage. This finding is consistent with the action of positive selection on amino-acid changes in the human lineage. However, the alternative hypothesis of a relaxation of constraints on *FOXP2* specific to the human lineage cannot be excluded on the basis of these data alone.

If these two changes in amino-acid encoding (or some other feature of the human *FOXP2* gene) were positively selected recently during human evolution, traces of a selective sweep should be detectable in the pattern of variation found among humans^{13,14}. To investigate this possibility, we sequenced a segment of 14,063 base pairs (bp) covering introns 4, 5 and 6 of the *FOXP2* gene in seven individuals from Africa, four from Europe, one from South America, five from mainland Asia and three from Australia and Papua New Guinea. In addition, we sequenced the same segment in a chimpanzee from central Africa, a chimpanzee from western Africa and an orang-utan (Table 1). One hallmark of a recent selective sweep is that more low-frequency alleles should be observed than expected under a neutral model of a random-mating population of constant size. To test this prediction, we calculated Tajima's D statistic¹⁵. The value is -2.20 for our sample, indicating a sharp excess of rare alleles. Under the standard neutral model outlined above, the probability of such an excess by chance is 0.002. Population growth can also lead to negative D values throughout the genome. However, the value of D at *FOXP2* is unusually low compared with other loci. For example, among 313 human genes¹⁶

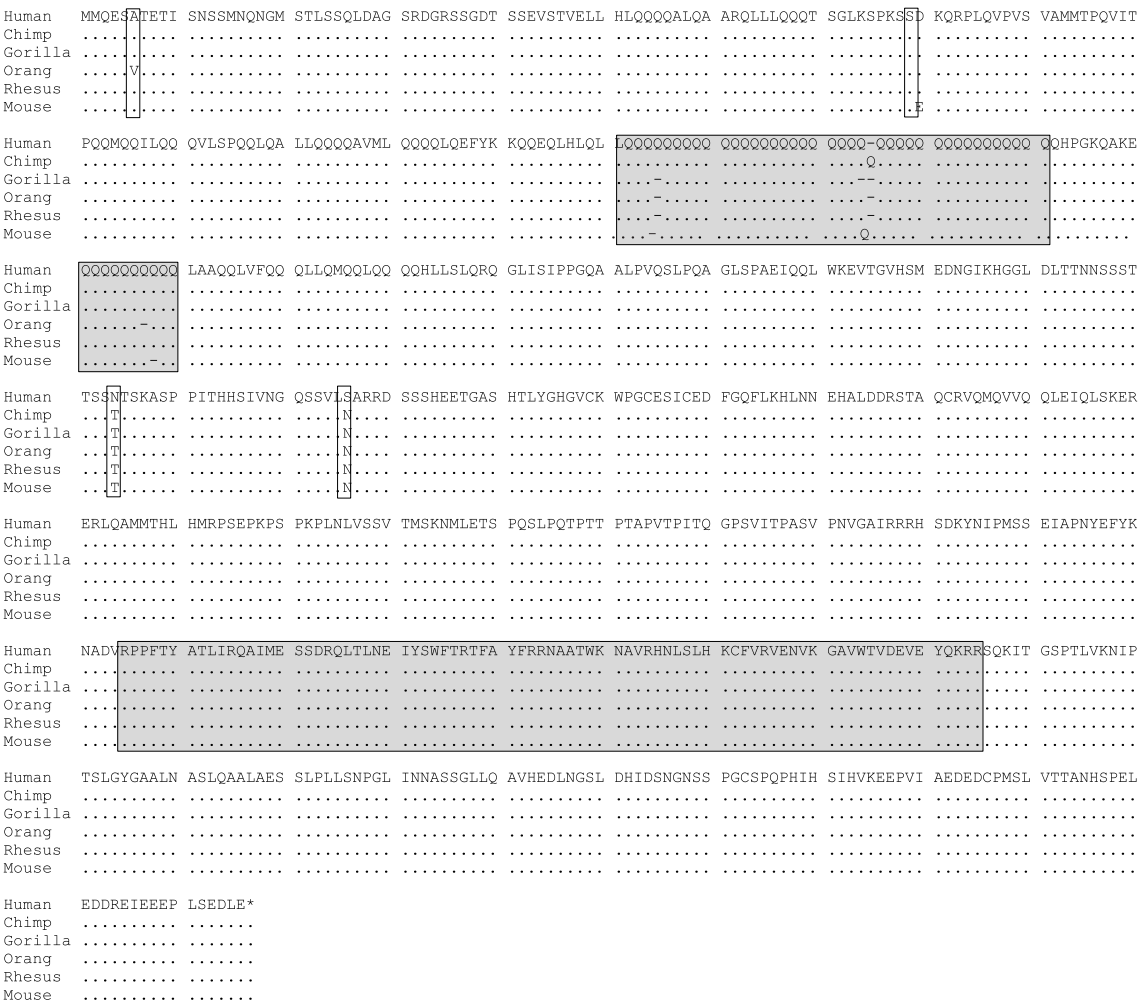


Figure 1 Alignment of the amino-acid sequences inferred from the *FOXP2* cDNA sequences. The polyglutamine stretches and the forkhead domain are shaded. Sites that differ from the human sequence are boxed.

sequenced in a sample of 164 chromosomes, only one has a more negative value (-2.25). A second prediction for a selective sweep at a recombining locus is that more derived (that is, non-ancestral) alleles at high frequency are expected than under the standard neutral model, a feature reflected in a negative H value¹⁷. To estimate H , we inferred the ancestral states of variable positions seen among the humans by using the chimpanzee and orang-utan DNA sequences. The H value of -12.24 deviates significantly from the neutral expectation of zero ($P = 0.042$) and would be even less likely by chance under a model with population growth¹³. The strongly negative D and H reflect an extreme skew in the frequency spectrum of allelic variants at *FOXP2* towards rare and high-frequency alleles. Because we considered a worldwide sample of humans, population structure might contribute to the negative D value. However, this type of sampling scheme is highly unlikely to produce a significantly negative H value. In contrast to demographic explanations, a selective sweep affecting the *FOXP2* gene can account for both aspects of the frequency spectrum. We do not observe a reduced diversity at human *FOXP2* relative to its divergence from the chimpanzee, as expected under a simple selective-sweep model. However, the magnitude of the reduction in variability expected after a selective sweep depends crucially on the rate of recombination. Estimates of recombination between intronic polymorphisms taken from a study of *FOXP2* (ref. 4) suggest that this region of the gene experiences rates of genetic exchange roughly five times the genome-wide average. If we assume that a selective sweep at a linked site does account for the patterns of variability recovered at *FOXP2*, it is noteworthy that the next gene is located 286 kilobases (kb) away from the sequenced segment. A selective sweep is not expected to lead to an excess of high-frequency derived alleles at sites that are 286 kb distant from the target of selection^{13,17}. Thus, the best candidates for the selected sites are the two amino-acid substitutions specific to humans in exon 7.

Individuals with disruption of *FOXP2* have multiple difficulties with both expressive and receptive aspects of language and grammar, and the nature of the core deficit remains a matter of debate^{18–20}. Nevertheless, a predominant feature of the phenotype of affected individuals is an impairment of selection and sequencing of fine orofacial movements¹⁸, an ability that is typical of humans and not present in the great apes. We speculate that some human-specific feature of *FOXP2*, perhaps one or both of the amino-acid substitutions in exon 7, affect a person's ability to control orofacial movements and thus to develop proficient spoken language. If this speculation is true, then the time when such a *FOXP2* variant became fixed in the human population may be pertinent with regard to the evolution of human language. We estimated this time point using a likelihood approach. Under a model of a randomly mating population of constant size, the most likely date since the fixation of the beneficial allele is 0, with approximate 95% confidence intervals of 0 and 120,000 years. Our point-estimate of 0

reflects the fact that high-frequency alleles rapidly drift to fixation, so an excess is most likely immediately after a selective sweep. However, if population growth soon succeeds the fixation of the advantageous allele, the rate of drift will be decreased and high-frequency alleles may persist longer in the population. Thus, the inclusion of population growth may push this time estimate back by at most the time since the onset of human population growth, some 10,000–100,000 years ago²¹. In any case, our method suggests that the fixation occurred during the last 200,000 years of human history, that is, concomitant with or subsequent to the emergence of anatomically modern humans²². This is compatible with a model in which the expansion of modern humans was driven by the appearance of a more-proficient spoken language²². However, to establish whether *FOXP2* is indeed involved in basic aspects of human culture, the normal functions of both the human and the chimpanzee *FOXP2* proteins need to be clarified. □

Methods

Isolation of cDNA sequences

For all analysed species, we amplified by polymerase chain reaction (PCR) and sequenced overlapping fragments of the *FOXP2* coding region from first-strand cDNA. Details are available in Supplementary Information.

Genomic sequencing

Full details are available in Supplementary Information. In brief, we designed primers from a human bacterial artificial chromosome (BAC) sequence (accession number AC020606), PCR-amplified fragments of 6–14 kb, re-amplified 2.2-kb fragments from these products that were then sequenced with internal primers. For each individual, each nucleotide position was read from both strands. Sequence traces were manually analysed for polymorphic positions using the program Seqman of the DNASTar package (see also Supplementary Information).

Data analysis

We aligned sequences with the help of the program ClustalW²³ and calculated most statistics with DnaSP 3.51 (ref. 24). P values for D and H were obtained by coalescent simulations implemented for a fixed number of segregating sites, and assuming no recombination. If we take into account recombination within the 14 kb, the P values decrease (for example, $P < 0.01$ for H and $P < 10^{-4}$ for D if one assumes an effective population size of 10^4 and a recombination rate of 5 centimorgans (cM) per Mb). Because the chimpanzee and orang-utan do not differ at any polymorphic position compared with humans, we assumed no back mutations when estimating the P value for H . The likelihood ratio tests for non-silent and silent substitutions were performed using the PAML package¹² as described²⁵ (see Supplementary Information). We predicted the structure of human, chimpanzee, mouse and orang-utan *FOXP2* using the program PredictProtein (<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>)⁶, which includes prediction of sites of protein kinase C phosphorylation by PROSITE²⁶. The orang-utan-specific alanine-to-valine change at position 6 results in the prediction of a β -sheet at positions 8–10 in the orang-utan, and the human-specific change at position 325 results in the prediction of a β -sheet in positions 323–326. However, these are not reliable and may not be relevant. We used the University of California at Santa Cruz Human Genome Project Working Draft, 22 December 2001 assembly (<http://genome.ucsc.edu>), to estimate distances to the closest genes. The middle of the sequenced region is 220 kb away from the known 5' end and 54 kb away from the 3' end of *FOXP2*, respectively. The next gene (supported by the cDNA sequence with GenBank accession number AF054589) is located 286 kb distant in the 3' direction.

Modelling the selective sweep

A summary likelihood method (compare with ref. 27) was used to estimate the time, T , since the fixation of the beneficial allele in the population. The polymorphism data was summarized as θ_H (ref. 17) and π (ref. 28). We then ran coalescent simulations of a selective sweep with recombination as in ref. 13. These simulations assume that we have polymorphism data for a neutral locus, at some distance from a selected site, and that selection acted on a newly arising variant. The likelihood of T is estimated as the proportion of n simulated data sets, where $|\theta_{\text{Hobs}} - \theta_{\text{Hsim}}| < \varepsilon$ and $|\pi_{\text{obs}} - \pi_{\text{sim}}| < \varepsilon$ (here, $n = 3 \times 10^6$ and $\varepsilon = 0.2$). The likelihood of T was evaluated over a grid of points spaced every 1,000 generations. We then chose the T value that maximizes the probability of obtaining the observed (θ_H , π) values. In addition to T , several additional parameters are in this selective sweep model: the distance to the selected site, the effective population size of humans, the strength of selection, the mutation rate and the recombination rate. It is not computationally feasible to co-estimate all of these parameters, and we proceeded by assuming that the values of most nuisance parameters are known exactly. We hypothesized that one of the substitutions on the human lineage was the selected site and used a point estimate of the population mutation rate (assuming 5 Myr to the common ancestor of a human and chimpanzee DNA sequence). We modelled uncertainty in the recombination rate per megabase by choosing the rate for each simulation from a γ distribution with parameters (5, 1); the mean was set to the recombination rate estimated from two polymorphic markers in introns 2 and 16, respectively, of the *FOXP2* gene⁴. The effective

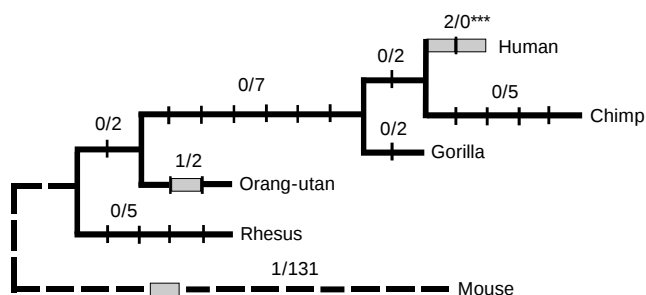


Figure 2 Silent and replacement nucleotide substitutions mapped on a phylogeny of primates. Bars represent nucleotide changes. Grey bars indicate amino-acid changes.

population size was taken to be 10^4 , on the basis of estimates for other loci²⁹. We tried three different values for the selection coefficient: $s = 5\%$, 1% and 0.5% . For these parameters, an s of 1% resulted in the highest likelihoods, so we reported the results for $s = 1\%$. If we use the chi-squared approximation with one degree of freedom for the log-likelihood ratio statistic $2\ln(\text{Lik}(\hat{T})/\text{Lik}(T))$, we obtain an approximate 95% confidence interval for T of $[0, 4,000 \text{ generations}]$. However, this approximation may not be appropriate in this context. Thus, we also ran 100 simulations to examine the distribution of T when the true T is equal to our maximum likelihood estimate of $T = 0$ (here, $n = 5 \times 10^5$ and $e = 0.2$). These simulations suggested an approximate 95% confidence interval of $[0, 6,000 \text{ generations}]$. We assumed a generation time of 20 years for converting T into years.

Received 11 November 2001; accepted 29 July 2002; doi:10.1038/nature01025.
Published online 14 August 2002.

- Liebermann, P. *The Biology and Evolution of Language* (Harvard Univ. Press, Cambridge, Massachusetts, 1984).
- Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).
- Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P. & Pembrey, M. E. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genet.* **18**, 168–170 (1998).
- Newbury, D. F. *et al.* *Foxp2* is not a major susceptibility gene for autism or specific language impairment. *Am. J. Hum. Genet.* **70**, 1318–1327 (2002).
- Makalowski, W. & Boguski, M. S. Evolutionary parameters of the transcribed mammalian genome: An analysis of 2,820 orthologous rodent and human sequences. *Proc. Natl Acad. Sci. USA* **95**, 9407–9412 (1998).
- Rost, B. Phd: Predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.* **266**, 525–539 (1996).
- Kops, G. J. *et al.* Control of cell cycle exit and entry by protein kinase b-regulated forkhead transcription factors. *Mol. Cell. Biol.* **22**, 2025–2036 (2002).
- Brunet, A. *et al.* Akt promotes cell survival by phosphorylating and inhibiting a forkhead transcription factor. *Cell* **96**, 857–868 (1999).
- Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).
- Eizirik, E., Murphy, W. J. & O'Brien, S. J. Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* **92**, 212–219 (2001).
- Chen, F. C. & Li, W. H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **68**, 444–456 (2001).
- Yang, Z. Paml: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
- Przeworski, M. The signature of positive selection at randomly chosen loci. *Genetics* **160**, 1179–1189 (2002).
- Simonsen, L., Churchill, G. A. & Aquadro, C. F. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**, 413–429 (1995).
- Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
- Stephens, J. C. *et al.* Haplotype variation and linkage disequilibrium in 313 human genes. *Science* **293**, 489–493 (2001).
- Fay, J. C. & Wu, C. I. Hitchhiking under positive darwinian selection. *Genetics* **155**, 1405–1413 (2000).
- Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. & Passingham, R. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl Acad. Sci. USA* **92**, 930–933 (1995).
- Gopnik, M. & Crago, M. B. Familial aggregation of a developmental language disorder. *Cognition* **39**, 1–50 (1991).
- Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain* **125**, 452–464 (2002).
- Wall, J. D. & Przeworski, M. When did the human population size start increasing? *Genetics* **155**, 1865–1874 (2000).
- Klein, G. *The Human Career, Human Biological and Cultural Origins* (Univ. Chicago Press, Chicago, 1989).
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
- Rozas, J. & Rozas, R. Dnasp version 3: An integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**, 174–175 (1999).
- Yang, Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.* **15**, 568–573 (1998).
- Bairoch, A., Bucher, P. & Hofmann, K. The prosite database, its status in 1997. *Nucleic Acids Res.* **25**, 217–221 (1997).
- Jensen, M. A., Charlesworth, B. & Kreitman, M. Patterns of genetic variation at a chromosome 4 locus of *Drosophila melanogaster* and *D. simulans*. *Genetics* **160**, 493–507 (2002).
- Tajima, F. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437–460 (1983).
- Harpending, H. & Rogers, A. Genetic perspectives on human origins and differentiation. *Annu. Rev. Genom. Hum. Genet.* **1**, 361–385 (2000).
- Watterson, G. A. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**, 256–276 (1975).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements

We thank F. Heissig for help with the cDNA sequencing; A. von Haeseler, G. Weiss and S. Zöllner for help with the data analysis on an earlier version of the manuscript; J. Wickings at the Centre International de Recherches Medicales for DNA samples of central chimpanzees; and the Bundesministerium für Bildung und Forschung, the Max

Planck Society and the Wellcome Trust for financial support. M.P. was supported by a National Science Foundation postdoctoral research fellowship in bioinformatics. S.E.F. is a Royal Society Research Fellow and A.P.M. is a Wellcome Trust Principal Research Fellow.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.P.

(e-mail: paabo@eva.mpg.de). *FOXP2* cDNA sequences of the mouse, rhesus macaque, orang-utan, gorilla, chimpanzee and human have GenBank accession numbers AY079003, AF512950, AF512949, AF512948, AF512947 and AF337817, respectively. Accession numbers for genomic sequences for the twenty humans, two chimpanzees and one orang-utan are AF515031–AF515050, AF515051–AF515052 and AF515053, respectively.

A corollary discharge maintains auditory sensitivity during sound production

James F. A. Poulet & Berthold Hedwig

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Speaking and singing present the auditory system of the caller with two fundamental problems: discriminating between self-generated and external auditory signals and preventing desensitization. In humans¹ and many other vertebrates^{2–7}, auditory neurons in the brain are inhibited during vocalization but little is known about the nature of the inhibition. Here we show, using intracellular recordings of auditory neurons in the singing cricket, that presynaptic inhibition of auditory afferents and postsynaptic inhibition of an identified auditory interneuron occur in phase with the song pattern. Presynaptic and postsynaptic inhibition persist in a fictively singing, isolated cricket central nervous system and are therefore the result of a corollary discharge from the singing motor network. Mimicking inhibition in the interneuron by injecting hyperpolarizing current suppresses its spiking response to a 100-dB sound pressure level (SPL) acoustic stimulus and maintains its response to subsequent, quieter stimuli. Inhibition by the corollary discharge reduces the neural response to self-generated sound and protects the cricket's auditory pathway from self-induced desensitization.

We have examined auditory information processing in the singing cricket, *Gryllus bimaculatus*. Males attract females or warn off rival males with songs that are generated by rubbing their forewings together. Calling song can be produced for many hours on end with a sound intensity greater than 100 dB SPL measured 50 mm from the ear⁸. It consists of a 250-ms series of chirps that is separated by a 300-ms chirp interval (Fig. 1a). A chirp itself comprises four syllables of 21-ms duration, which are generated by the closing movements of the forewings. Crickets' ears are located on their forelegs and are therefore exposed fully to the self-generated sounds. About 60 auditory afferent neurons project from the ear along the fifth prothoracic nerve and terminate in the auditory neuropile of the prothoracic ganglion⁹. Two local, mutually inhibitory omega 1 neurons (ON1s) are located in the prothoracic ganglion¹⁰. This identified pair of bilaterally symmetrical interneurons receives auditory information from the ear ipsilaterally to their soma. These interneurons are most sensitive to the carrier frequency (4.5 kHz) of the male calling song.

Crickets must maintain auditory sensitivity during bouts of singing because they respond behaviourally to auditory stimulation

during chirp intervals¹¹. They do not, unlike some other animals^{12–15}, modulate the responsiveness of their peripheral auditory system during sound production¹⁶. To understand how auditory sensitivity is maintained despite the intense self-stimulation, we examined the activity of the auditory afferents and ON1 to acoustic stimulation during sonorous, silent and fictive singing.

During pharmacologically elicited sonorous singing, we made intracellular recordings from the dendritic region of ON1. The neuron responded with bursts of spikes in phase with the syllables (Fig. 1a). The maximum spike frequency of ON1 to self-generated syllables was on average 176 ± 28 Hz (mean $\pm$ s.e.m; $n = 10$ crickets). As this was much lower than the response of ON1 to 100 dB SPL sound pulses at rest (376 ± 50 Hz, $n = 5$ crickets), it indicated that an inhibitory input influenced auditory processing during singing.

If singing was elicited but sound production was prevented by removing one forewing to cause 'silent singing' then ON1 did not spike, which indicated that its response during sonorous singing was caused by the crickets' song. Instead, ON1 received inhibitory postsynaptic potentials (IPSPs) during the syllables (Fig. 1b). In five silently singing crickets, the average duration of a wing movement cycle underlying sound production was 38.1 ± 1.5 ms. IPSPs started 5.0 ± 0.2 ms after the start of wing closing and reached a maximum at 23.3 ± 1.4 ms, just after the transition between wing closing and opening at 19.0 ± 0.7 ms. If a sequence of acoustic stimuli was presented during silent singing, then ON1 spiked continuously during the chirp intervals but was inhibited during the chirps (Fig. 1c). Only excitatory postsynaptic potentials (EPSPs), and occasionally a spike, were elicited by the stimuli

during silent chirps. Thus, the activity of ON1 during sonorous stridulation is due to excitation from refferent sound stimulation, which is caused by the crickets' own song and a synchronous inhibitory input.

To determine whether this inhibition was elicited by sensory feedback or by the network of neurons that generate the motor pattern, we induced fictive singing in crickets with either their thoracic ($n = 3$) or thoracic and abdominal ganglia ($n = 2$) isolated from muscles and sense organs, except for the fifth prothoracic nerve. IPSPs were also present in the fictively singing cricket. The IPSPs were sufficient to suppress the spiking response to acoustic stimuli (Fig. 1d), as in silent singing. The IPSPs even persisted when the ears were removed ($n = 4$; Fig. 1e).

To analyse whether auditory afferents are also affected by inhibition, we made intracellular recordings close to their terminals. During sonorous singing, the afferents responded in phase with sound production (Fig. 2a). During quieter chirps, the presence of primary afferent depolarizations (PADs) before the spikes was indicated by a gradual depolarization (Fig. 2a, arrow). The PADs might be masked by the spike activity, we therefore recorded auditory afferents in six silently singing crickets. Spikes were only occasionally produced during silent chirps, which confirmed that the afferents responded to the self-generated sound. PADs were now obvious during the syllables (Fig. 2b).

In four silently singing crickets, the average duration of a wing movement cycle was 36.0 ± 1.8 ms. The PADs started 4.2 ± 0.4 ms after the start of wing closing and reached a maximum at 19.3 ± 1.4 ms, just after the transition between wing closing and opening at 17.8 ± 1.0 ms. Thus, the PADs have a similar timing to

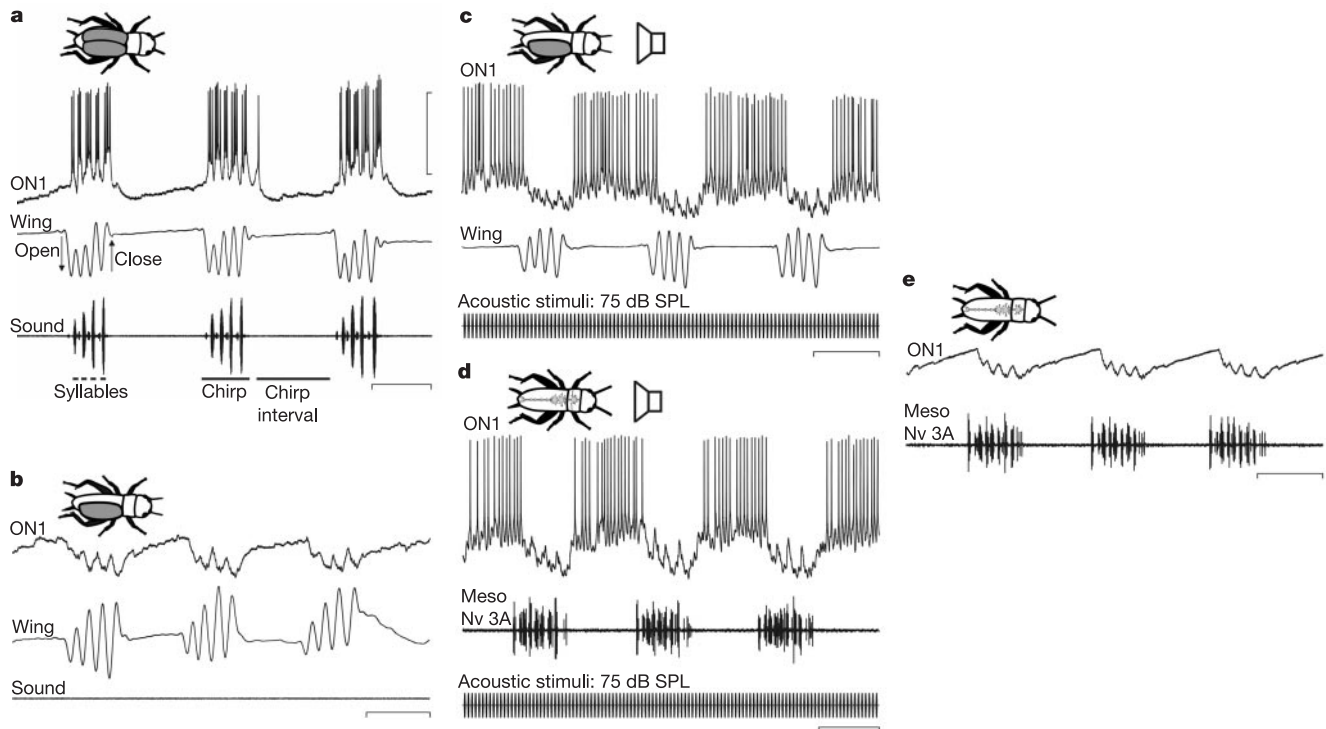


Figure 1 Responses of ON1 during singing. **a**, During sonorous singing ON1 produces bursts of spikes in phase with each syllable. A syllable, a whole chirp and a chirp interval are marked below the sound recording. **b**, During silent singing, IPSPs are present in ON1 in phase with the silent wing movements. They increase in amplitude from -2.3 ± 0.3 mV at the first closing wing movement to -5.3 ± 0.5 mV at the last ($n = 8$ crickets). **c, d**, When stimulated with a sequence of acoustic pulses, ON1 responds during the chirp intervals, but is inhibited during both silent (**c**) and fictive (**d**) chirps. **e**, During fictive singing, IPSPs occur in phase with the chirps even when the animal's ears are

removed. Mean amplitude of the IPSPs increases from -2.4 ± 0.6 mV for the first syllable of the fictive chirp to -5.1 ± 0.7 mV at the last ($n = 5$ crickets). Symbols above the figures represent singing with two wings, singing with one wing, fictively singing and fictively singing with ears removed. ON1, intracellular ON1 recording; wing, wing movements; sound, cricket song; acoustic stimuli, sound pulses; Meso Nv 3A, activity of mesothoracic nerve 3A. Vertical scale bar: intracellular, 25 mV; extracellular, 10 mV; wing, 1 mm. Horizontal scale bars, 250 ms.

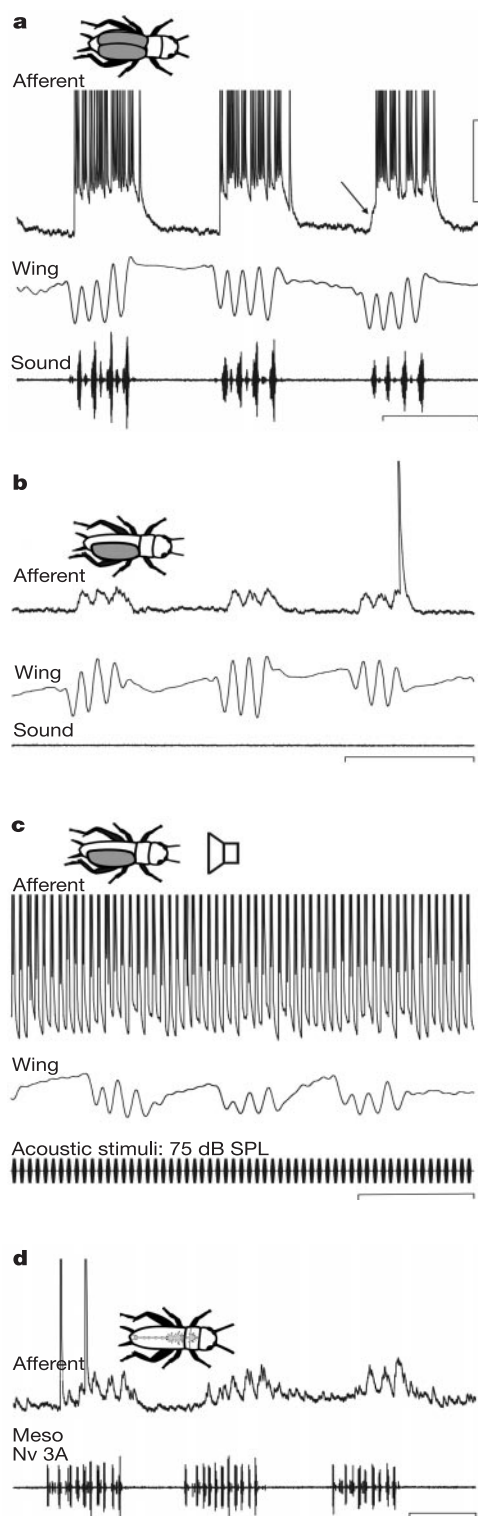


Figure 2 Presynaptic inhibition of auditory afferents during singing. **a**, Auditory afferents respond in phase with sound production. Arrow indicates a primary afferent depolarization (PAD) at the start of a chirp. **b**, During silent singing, PADs are present in auditory afferents during the rhythmic wing movements. They increase in amplitude from $+2.6 \pm 0.5$ mV at the start of the chirp to $+3.3 \pm 0.7$ mV at the end ($n = 6$ crickets). **c**, The spike pattern of an auditory afferent presented with a sequence of 4.5 kHz, 75 dB SPL sound pulses is not modulated during silent singing. **d**, PADs are also present during fictive chirps which increase in amplitude from 1.9 ± 0.3 mV at the start of the fictive chirp to 3.5 ± 0.1 mV at the end ($n = 6$ crickets). Afferent, intracellular recording of an auditory afferent, the spikes have been truncated. Vertical scale bar: intracellular, 15 mV; extracellular, 10 mV; wing, 1 mm (**a**), 0.25 mm (**b**), 0.5 mm (**c**). Horizontal scale bars, 250 ms.

the IPSPs in ON1. The PADs did not affect spike production in the auditory afferents, as they spiked consistently to acoustic stimulation throughout silent singing (Fig. 2c). PADs were also present in six fictively singing crickets (Fig. 2d), even when their ears were removed ($n = 2$ crickets). In many sensory systems, PADs have an inhibitory function¹⁷ as they reduce synaptic efficacy at the afferent terminals. The ultrastructural and histochemical prerequisites for presynaptic inhibition have been described in the cricket¹⁸.

We tested the effectiveness of the inhibition mediated by the corollary discharge (sum of presynaptic and postsynaptic inhibition) on auditory processing in ON1 by playing acoustic stimuli that were similar to the natural calling song during silent singing (Fig. 3a). At rest and during the chirp intervals, each sound pulse evoked a depolarization and a burst of spikes in ON1 with an average maximum spike frequency of 376 ± 50 Hz ($n = 5$ crickets). During silent chirps ON1 responded to the stimuli with bursts of only 123 ± 29 Hz, which was significantly lower than the response during the intervals (two-tailed paired t -test: $P < 0.002$, $t = 7.71$, degrees of freedom (d.f.) = 4; Fig. 3b).

We then tested the effect of the inhibition on the sensitivity of ON1. Intense acoustic stimulation causes ON1 to spike, but it also causes a graded hyperpolarization that reduces its sensitivity to sound¹⁹. The hyperpolarization is thought to be due to Ca^{2+} -dependent K^{+} channels that are activated by the spikes²⁰. The response to self-generated sound should therefore desensitize the response of ON1 to external sounds in the absence of a corollary discharge. We mimicked the effects of the animal listening to its own

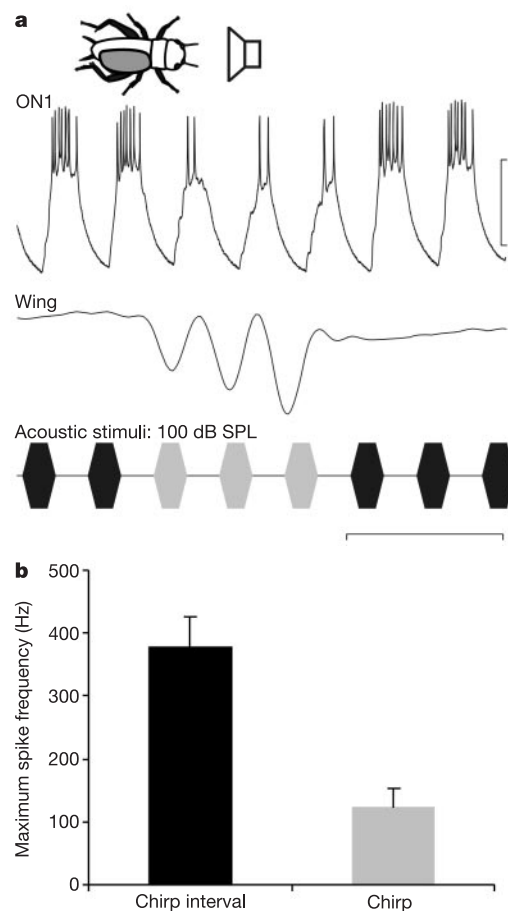


Figure 3 Testing the efficacy of inhibition during silent singing. **a**, A typical ON1 response to 100 dB SPL sound pulses during the chirp interval (in black) and during a silent chirp (in grey). **b**, Averaged maximum spike frequency to 150 sound pulses presented to five silently singing crickets during the chirp interval (in black) and the chirp (in grey). Vertical scale bar: intracellular, 25 mV; wing, 1 mm. Horizontal scale bar, 100 ms.

song and examined its responses to quieter test stimuli ($n = 8$ crickets). First a series of test sound pulses was presented at 80 dB SPL that each elicited a burst of spikes with an average maximum frequency of 203 ± 24 Hz (Fig. 4a). When normal singing was mimicked with 100 dB SPL chirps, the response to the following 80 dB SPL test stimuli was reduced so that very few spikes were elicited (Fig. 4b) with an average maximum spike frequency of

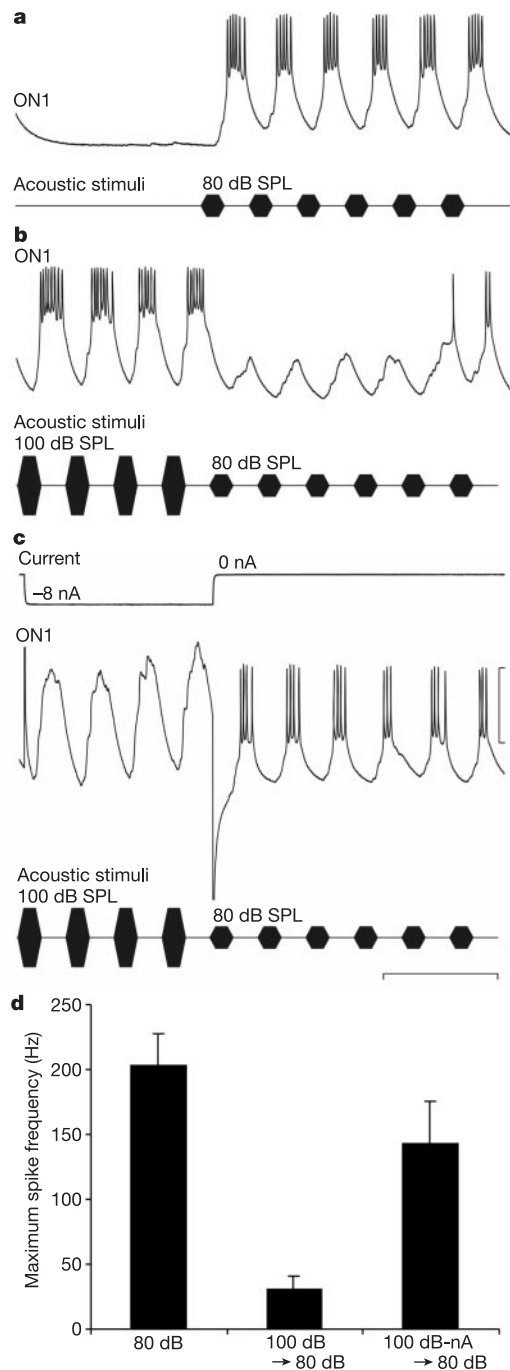


Figure 4 Inhibition of ON1 during acoustic stimulation prevents subsequent desensitization. **a**, In a control experiment, ON1 produces bursts of spikes at 203 Hz in response to 80 dB SPL pulses. **b**, Response of ON1 to 80 dB SPL sound pulses is reduced to 30 Hz if the sound pulses are preceded by a 100 dB SPL chirp. **c**, If spiking is prevented by hyperpolarizing current injection during responses to 100 dB SPL chirps, the response to the subsequent 80 dB SPL pulses is 143 Hz. **d**, Pooled, averaged maximum spike frequency of eight crickets to 150 consecutive presentations of each stimulus paradigm. Current, current injected into the cell. Vertical scale bar: intracellular, 25 mV; current, 20 nA. Horizontal scale bar, 100 ms.

30 ± 11 Hz. This was significantly different to the response in the control situation (two-tailed paired t -test: $P < 0.001$, $t = 8.70$, d.f. = 7; Fig. 4a, b, d). The reduction was greatest just after the 100 dB SPL chirp but persisted for the rest of the artificial chirp interval.

We then injected hyperpolarizing current into ON1 to prevent it from spiking during the 100 dB SPL chirp. The average maximum response to subsequent 80 dB SPL stimuli was a burst of spikes at 143 ± 32 Hz, which was significantly higher than the response without current injection (two-tailed paired t -test, $P < 0.004$, $t = -4.21$, d.f. = 7; Fig. 4a, c, d). The response after current injection did not reach the response to the test stimuli, because afferent adaptation occurs at high sound intensities²¹. This effect was still present after deafening the ear contralateral to the soma of ON1 and was therefore independent from inhibition from the contralateral ON1. In three one-eared crickets, the average spike frequency to 80 dB SPL stimuli was 263 Hz; this was reduced to 23 Hz when the stimuli were preceded by a 100 dB SPL chirp but increased to 168 Hz if the neuron was hyperpolarized during the chirp. Thus, a reduction in the spiking response of ON1 to intense sounds maintains the neuron's sensitivity to subsequent sounds.

To allow animals to distinguish between reafferent and external stimuli, the two types of sensory information must be treated differently. It was proposed 50 years ago that efferent neural signals, efference copies²² or corollary discharges²³, from central motor networks alter the responsiveness of sensory pathways in phase with the generation of reafferent sensory information. Both mechanisms could prevent self-induced desensitization and help discriminate important features from the mix of self-generated and external sensory information. These mechanisms have now been identified in several sensory systems^{24–27}. In auditory systems, inhibition of auditory neurons has been seen in many vertebrates^{1–7}, but intracellular recordings have never been obtained during sound production to establish the exact site and mode of the inhibition.

We have shown here that in the cricket an efferent signal modulates auditory information processing at two levels of the auditory system: it causes PADs in auditory afferent terminals and IPSPs in ON1. This efferent signal may be termed a 'corollary discharge' because it is generated in the nervous system and the strength of its inhibition is independent of sound production²⁸. The neuron(s) that mediate the corollary discharge are unlikely to be the descending command neurons for stridulation, as they spike with little modulation of their firing rates during singing²⁹. The most likely candidate neurons are those contained in the thoracic stridulatory motor network. The inhibition significantly reduces the response of ON1 during sound production and thereby maintains this neuron's sensitivity to external sounds in the chirp intervals. In this way, the corollary discharge prevents auditory desensitization and allows crickets to respond to auditory signals during singing. □

Methods

Crickets

We carried out experiments at 18–22 °C on adult male *Gryllus bimaculatus* from a colony maintained on a 12 h/12 h light/dark cycle. Crickets were fixed on a holder in a standing position to allow free wing movement. The holder was then rotated 180° for intracellular recording from prothoracic neurons and pharmacological stimulation of the brain. For silent singing the left wing was removed. Exposed nervous tissue was bathed in insect saline (ionic composition (mM): NaCl, 140; KCl, 10; CaCl₂, 4; NaHCO₃, 4; NaHPO₄, 6). The acetylcholine esterase inhibitor eserine salicylate (10^{-2} M) was pressure-injected into the anterior protocerebrum of the brain through a glass microcapillary to elicit singing.

Acoustic stimuli

Acoustic stimuli were generated by the software Cool Edit 2000 and had a frequency of 4.5 kHz with rising and falling ramps of 2 ms. Acoustic stimuli that mimicked natural singing had a duration of 21 ms, 42-ms intervals and were 100 dB SPL relative to 20 μ Pa root mean square (r.m.s.). To reveal the time course of the neuronal response during singing, we presented a sequence of 75-dB SPL (r.m.s.) acoustic stimuli with a short interval (15 ms) and duration (8 ms). Sound amplitude was calibrated with a Brüel and Kjær measuring amplifier (type 2610) and a microphone (type 4191) positioned at the cricket's ears.

Data recording and neuron visualization

Auditory neurons were recorded and stained using thick-walled glass micropipettes filled with 5% Lucifer yellow and 0.5 M LiCl (resistance 100–150 MΩ). After recording, ganglia were processed conventionally and the stained neurons were identified under an ultraviolet fluorescence microscope. To record fictive motor activity, we placed a suction electrode on mesothoracic nerve 3A, which contains motor axons that innervate wing closer and opener muscles. A microphone (Audio-Technica AT853A) recorded sound produced by the cricket and an optoelectronic camera monitored wing movements. All data were transferred directly onto a computer through an AD board (Data Translation 2821 F8DI) with a sampling rate of 10 kHz per channel. We analysed data off-line using NeuroLab³⁰ and Microsoft Excel 2000.

Received 19 February; accepted 23 May 2002; doi:10.1038/nature00919.

- Creutzfeldt, O., Ojemann, G. & Lettich, E. Neuronal activity in the human temporal lobe II. Responses to the subjects own voice. *Exp. Brain Res.* **77**, 476–489 (1989).
- Suga, N. & Schlegel, P. Neural attenuation of responses to emitted sounds in echolocating bats. *Science* **177**, 82–84 (1972).
- Suga, N. & Shimozawa, T. Site of neural attenuation of responses to self-vocalized sounds in echolocating bats. *Science* **183**, 1211–1213 (1974).
- Schuller, G. Vocalization influences auditory processing in collicular neurons of the CF-FM bat, *Rhinolophus ferrumequinum*. *J. Comp. Physiol. A* **132**, 39–46 (1979).
- McCasland, J. S. & Konishi, M. Interaction between auditory and motor activities in an avian song control nucleus. *Proc. Natl Acad. Sci. USA* **78**, 7815–7819 (1981).
- Müller-Press, P. & Ploog, D. Inhibition of auditory cortical neurons during phonation. *Brain Res.* **215**, 61–76 (1981).
- Metzner, W. A possible neuronal basis for Doppler-shift compensation in echo-locating horseshoe bats. *Nature* **341**, 529–532 (1989).
- Nocke, H. Physiological aspects of sound communication in crickets (*Gryllus campestris* L.). *J. Comp. Physiol. A* **80**, 141–162 (1972).
- Michel, K. Das Tympanalorgan von *Gryllus bimaculatus* DeGeer (Saltatoria, Gryllidae). *Z. Morph. Tiere* **77**, 285–315 (1974).
- Schildberger, K., Wohlers, D. W. & Huber, F. in *Cricket Behaviour and Neurobiology* (eds Huber, F., Moore, T. E. & Loher, T. E.) 423–458 (Cornell Univ. Press, Ithaca/London, 1989).
- Jones, M. D. R. & Dambach, M. Response to sound in crickets without tympanal organs (*Gryllus campestris* L.). *J. Comp. Physiol. A* **87**, 89–98 (1973).
- Suga, N. & Jen, P. Peripheral control of acoustic signals in the auditory system of echolocating bats. *J. Exp. Biol.* **62**, 277–311 (1975).
- Borg, E. & Counter, S. The middle-ear muscles. *Sci. Am.* **261** (August), 62–68 (1989).
- Narins, P. M. Reduction of tympanic membrane displacement during vocalization of the arboreal tree frog, *Eleutherodactylus coqui*. *J. Acoust. Soc. Am.* **91**, 3551–3557 (1992).
- Hennig, R. M. et al. Auditory threshold change in singing cicadas. *J. Exp. Biol.* **187**, 45–55 (1994).
- Poulet, J. F. A. & Hedwig, B. Tympanic membrane oscillations and auditory receptor activity in the stridulating cricket *Gryllus bimaculatus*. *J. Exp. Biol.* **204**, 1281–1293 (2001).
- Clarac, F. & Cattarct, D. Invertebrate presynaptic inhibition and motor control. *Exp. Brain Res.* **112**, 163–180 (1996).
- Hardt, M. & Watson, A. H. D. Distribution of input and output synapses on the central branches of bushcricket and cricket auditory afferent neurones: immunocytochemical evidence for GABA and glutamate in different populations of presynaptic boutons. *J. Comp. Neurol.* **403**, 281–294 (1999).
- Pollack, G. S. Selective attention in an insect auditory neuron. *J. Neurosci.* **8**, 2635–2639 (1988).
- Sobel, E. C. & Tank, D. W. *In vivo* Ca²⁺ dynamics in a cricket auditory neuron: an example of chemical computation. *Science* **263**, 823–826 (1994).
- Givois, V. & Pollack, G. S. Sensory habituation of auditory receptor neurons: implications for sound localization. *J. Exp. Biol.* **203**, 2529–2537 (2000).
- von Holst, E. & Mittelstaedt, H. Das reafferenzprinzip. (Wechselwirkungen zwischen zentralnervensystem und peripherie.). *Naturwissenschaften* **37**, 464–476 (1950).
- Sperry, R. W. Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psych.* **43**, 482–489 (1950).
- Zaretsky, M. & Rowell, C. H. F. Saccadic suppression by corollary discharge in the locust. *Nature* **280**, 583–585 (1979).
- Bell, C. C. An efference copy which is modified by reafferent input. *Science* **214**, 450–453 (1981).
- Guthrie, B. L., Porter, J. D. & Sparks, D. L. Corollary discharge provides accurate eye position information to the oculomotor system. *Science* **221**, 1193–1195 (1983).
- Sillar, K. T. & Roberts, A. A neuronal mechanism for sensory gating during locomotion in a vertebrate. *Nature* **331**, 262–265 (1988).
- Bell, C. C. in *Comparative Physiology of Sensory Systems* (eds Bolis, L., Keynes, R. D. & Maddrell, S. H. P.) 636–647 (Cambridge Univ. Press, Cambridge, 1984).
- Hedwig, B. Control of cricket stridulation by a command neuron: efficacy depends on the behavioural state. *J. Neurophysiol.* **83**, 712–722 (2000).
- Knepper, M. & Hedwig, B. NEUROLAB, a PC-program for the processing of neurobiological data. *Comp. Methods Programs Biomed.* **52**, 75–77 (1997).

Acknowledgements

We thank M. Burrows, T. Matheson and S. Rogers for comments on the manuscript. This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) studentship and grants from the Wellcome Trust and the Royal Society.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.P. (e-mail: jfap2@cam.ac.uk) or B.H. (e-mail: bh202@cam.ac.uk)

Mechanism of magnesium activation of calcium-activated potassium channels

Jingyi Shi*, Gayathri Krishnamoorthy*, Yanwu Yang†, Lei Hu*, Neha Chaturvedi*, Dina Harilal*, Jun Qin† & Jianmin Cui*

* Cardiac Bioelectricity Research and Training Center and Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio 44106-7207, USA

† Structural Biology Program, Lerner Research Institute, The Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA

Large-conductance (BK type) Ca²⁺-dependent K⁺ channels are essential for modulating muscle contraction and neuronal activities such as synaptic transmission and hearing^{1–5}. BK channels are activated by membrane depolarization and intracellular Ca²⁺ and Mg²⁺ (refs 6–10). The energy provided by voltage, Ca²⁺ and Mg²⁺ binding are additive in activating the channel, suggesting that these signals open the activation gate through independent pathways^{9,11}. Here we report a molecular investigation of a Mg²⁺-dependent activation mechanism. Using a combined site-directed mutagenesis and structural analysis, we demonstrate that a structurally new Mg²⁺-binding site in the RCK/Rossman fold domain—an intracellular structural motif that immediately follows the activation gate S6 helix^{12–15}—is responsible for Mg²⁺-dependent activation. Mutations that impair or abolish Mg²⁺ sensitivity do not affect Ca²⁺ sensitivity, and vice versa. These results indicate distinct structural pathways for Mg²⁺- and Ca²⁺-dependent activation and suggest a possible mechanism for the coupling between Mg²⁺ binding and channel opening.

The energetically separate Ca²⁺- and Mg²⁺-dependent activation pathways suggest that each pathway may involve distinct structural components of the channel. Previous results suggest that a low-affinity, divalent cation-binding site that is responsible for Mg²⁺-dependent activation may be located in the amino-terminal core of mouse Slo1 (mSlo1) subunits^{9,16} (Fig. 1a). Thus, the Mg²⁺-binding site is distinct from the high-affinity Ca²⁺-binding site that has been proposed to reside in the carboxy terminal tail^{17–19} (Fig. 1a). The RCK domain is an intracellular motif of the core that immediately follows the activation gate S6 helix^{12–15} (Fig. 1a). It is conserved among BK channels, various prokaryotic K⁺ channels and TrkA proteins (which regulate K⁺ conductance¹²). The X-ray crystal structure of the RCK domain of the *Escherichia coli* Kch channel indicates that this domain may contain a ligand-binding site at its N-terminal half². To examine whether the N terminus of the mSlo1 RCK domain contains the low-affinity metal-binding site, we first studied chimaeric channels between mSlo1 and its homologue mSlo3 (ref. 20)—the activation of which is insensitive to Mg²⁺, although it also contains the RCK domain^{9,12} (Fig. 1). Comparing the sequence of mSlo1 with mSlo3, it is obvious that differences scatter within the N-terminal region of the RCK domain (Fig. 2a). If these differences occur in the metal-binding site, they may result in the difference in Mg²⁺ sensitivity between these two channels. Figure 1c (left panel) shows that the conductance–voltage (G–V) relation of chimaera C31-I (see Methods for definition of chimaeras) shifted less than that of mSlo1 when intracellular Mg²⁺ concentration ([Mg²⁺]_i) increased from 0 to 10 mM, whereas the increase of [Mg²⁺]_i caused no change in the G–V relation of C31-II. The total loss of Mg²⁺ sensitivity in C31-II is consistent with the idea that the sequence of mSlo3 in this region may have destroyed the metal-binding site. On the other hand, chimaera C13 was activated by Mg²⁺ (Fig. 1c, right panel), indicating that this region in mSlo1 is sufficient to restore Mg²⁺ sensitivity in mSlo3. Figure 1d

compares the free energy provided by Mg^{2+} binding towards the activation of mSlo1, mSlo3 and chimaeric channels^{9,11} (see Methods and Supplementary Information), and quantitatively demonstrates that the mSlo1 sequence at the N terminus of the RCK domain is essential for Mg^{2+} -dependent activation.

To identify individual amino acids that are important for Mg^{2+} -dependent activation we studied the effects of site-directed mutations on the Mg^{2+} sensitivity of mSlo1 channels. The core structure of the RCK domain adopts a Rossmann fold with a six-stranded parallel β -sheet (β A– β F) and α -helices (α A– α E) on both sides¹². Figure 2a shows the sequence alignment of the N terminus of the RCK domain that includes β A– α C in mSlo1, *Drosophila* Slo (dSlo)²¹, mSlo3 and *E. coli* Kch²² channels. The activation of dSlo channels has a similar Mg^{2+} sensitivity as that of mSlo1 (data not shown). Residues that are conserved in mSlo1 and dSlo but not in mSlo3 (Fig. 2a) are probably responsible for the difference in Mg^{2+} sensitivities among these channels. Therefore, in one series of mutations these residues in mSlo1 were mutated into the corresponding amino acids in mSlo3, either individually or in combination (Fig. 2b). In another series of mutations, oxygen-containing residues that are conserved in mSlo1 and dSlo—which may possibly be Mg^{2+} -coordinating—were mutated into Ala or other amino acids as indicated in Fig. 2b. The effects of these mutations on Mg^{2+}

sensitivity show a clear pattern with regard to the position of mutated residues (Fig. 2b). Mutations at the interloop connecting β A and α A (H350N), both N- and C-termini of β B (E374A, H379G), and towards the N terminus of β C (T396A, Q397C and E399N) reduced Mg^{2+} sensitivity, whereas mutations at other segments had no significant effects on Mg^{2+} sensitivity. Most notably, a change in either of two residues, E374A or E399N, completely abolished Mg^{2+} -dependent activation at 110 μ M $[Ca^{2+}]_i$ (Fig. 2c). The large effect of the mutations E374A and E399N on Mg^{2+} sensitivity was not accompanied by any change in Ca^{2+} -dependent activation or activation in the absence of Ca^{2+} and Mg^{2+} (Fig. 2c), suggesting that no gross change of channel structure had resulted from these mutations. E374 is conserved among mSlo1, dSlo and mSlo3, whereas E399 is conserved in mSlo1 and dSlo but not in mSlo3 (Fig. 2a), which explains why the chimaera

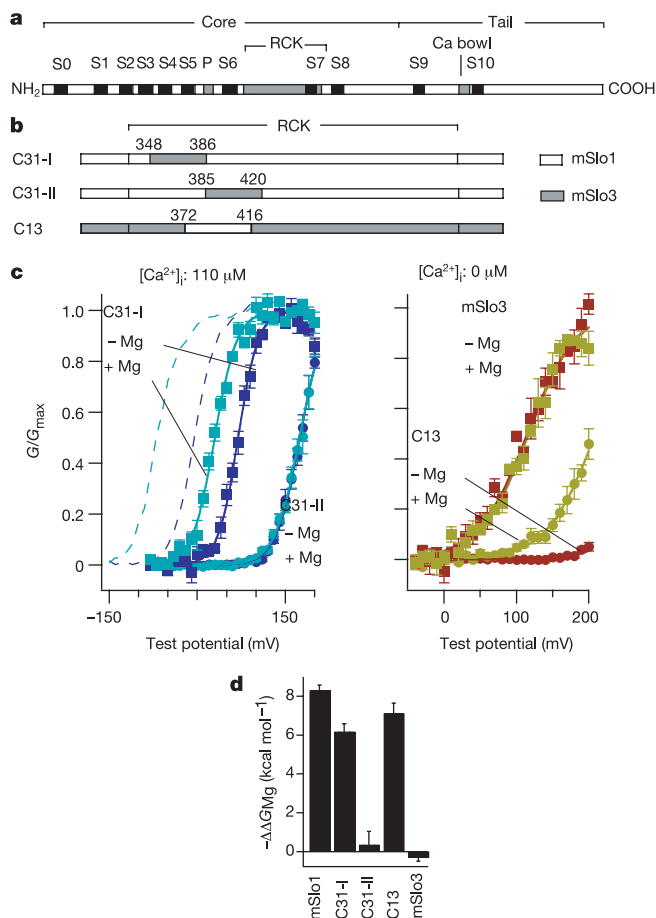


Figure 1 Results of chimaeric channels. **a**, The Slo polypeptide. S0–S6, transmembrane segments; P, pore loop; S7–S10, cytoplasmic hydrophobic segments; Ca bowl, putative high-affinity Ca^{2+} -binding site¹⁸. **b**, Constructs of chimaeric channels C31-I, C31-II and C13. Numbers indicate the position in mSlo1 where the substitution starts and ends. **c**, Mean G - V relations of C31-I, C31-II, mSlo3 and C13 at 0 and 10 mM $[Mg^{2+}]_i$ ($n = 5$ –6 patches). Dashed lines are G - V relations of mSlo1 for comparison (Supplementary Information). **d**, Free energy provided by Mg^{2+} binding towards the activation of channels when $[Mg^{2+}]_i$ increases from 0 to 10 mM.

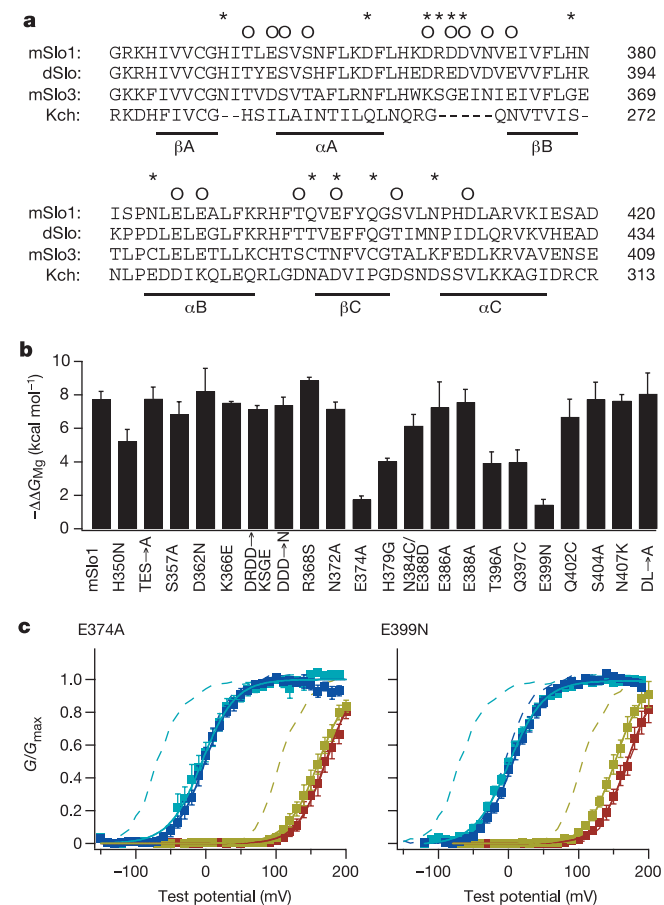


Figure 2 Results of site-directed mutations. **a**, Sequence alignment of part of the RCK domain in mSlo1, dSlo, mSlo3 and *E. coli* Kch channels¹². Underlines indicate amino acids that form α -helices (α A– α C) and β -strands (β A– β C). Numbers indicate the position of the far right residues in the primary sequence of their respective proteins. Asterisk, residues that are conserved in mSlo1 and dSlo, but not in mSlo3; O, oxygen-containing residues that are conserved in mSlo1 and dSlo. **b**, Free energy provided by Mg^{2+} binding in activating the wild type and mutant mSlo1 channels when $[Mg^{2+}]_i$ increases from 0 to 10 mM. Free energy was measured at 0 $[Ca^{2+}]_i$, except for TES $\rightarrow$ A and H379G (measured at 110 μ M $[Ca^{2+}]_i$). For single mutations, the original amino acid of mSlo1, its position, and the amino acid that it is changed to are indicated. TES $\rightarrow$ A, T352, E354 and S355 are changed to A; DRDD $\rightarrow$ KSGE, DRDD (positions 367–370) are changed to KSGE; DDD $\rightarrow$ N, D367, D369 and D370 are changed to N; N384C/E388D, a double mutant N384C and E388D; DL $\rightarrow$ A, D410 and L411 are changed to A. $n = 3$ –6 patches for each channel. **c**, Mean G - V relations of E374A and E399N channels. $[Mg^{2+}]_i$ and $[Ca^{2+}]_i$ are colour-coded as in Fig. 1c. Dashed lines are G - V relations of mSlo1 for comparison (Supplementary Information).

C31-II lost Mg^{2+} -sensitivity but C31-I did not (Fig. 1). These results suggest that residues E374 and E399 are part of the low-affinity, metal-binding site, which are exposed to the aqueous solution and are coordinated to the bound Mg^{2+} ion. Therefore, mutation of these residues should have little impact on the channel structure, but should significantly affect Mg^{2+} binding.

Figure 3 shows the structure of the N terminus of the RCK domain of *E. coli* Kch channels¹² in which the residues at positions corresponding to E374, Q397, E399 and H379 of mSlo1 (N267, N289, D292 and S272 in Kch; Fig. 2a) are substituted by mSlo1 counterparts. A metal-binding site with a bound Mg^{2+} ion is shown as the result of energy minimization. The Mg^{2+} ion is coordinated by the side chains of E374 and E399 (located at the N terminus of two adjacent β -strands, β B and β C), and Q397 (in the loop connecting the N terminus of β C with α B). The main chain carbonyl oxygen of Q397 also coordinates to the Mg^{2+} ion (Fig. 3). These results are consistent with mutational results (Figs 1 and 2). However, the location of the site was unexpected because in other proteins with a Rossman fold—such as the response regulator of bacterial chemotaxis, CheY²³, and integrin domain A²⁴—the ligand-binding site is always formed by loops connecting the C terminus of the parallel β -strands to helices²⁵ (indicated by the arrow in Fig. 3). Nevertheless, such an organization of metal coordination in mSlo1 is comparable to that found in CheY, in which three Asp side chains and one backbone carbonyl are coordinated to the Mg^{2+} ion²³. The metal site in CheY selectively binds divalent or trivalent cations but does not discriminate between metals on the basis of their size²⁶. Similarly, the low-affinity metal site in mSlo1 channels binds Mg^{2+} with a comparable affinity as Ca^{2+} at the millimolar range^{9,10}, which is around the physiological intracellular Mg^{2+} concentration²⁷. Therefore, in a typical cell, this site is essentially a Mg^{2+} -binding site.

The results of structural analysis reveal that Q397 is one of the Mg^{2+} -coordinating residues. The mutation Q397C reduced Mg^{2+} sensitivity of the channel but did not completely abolish it at 0 or 110 μ M $[Ca^{2+}]_i$ (Fig. 4a). Thus, the coordination by the side chain of Q397 is probably replaced either by a water molecule or the thiol group of Cys, and thus maintains some Mg^{2+} sensitivity. When an Asp residue, which is also negatively charged but with a shorter side

chain, replaced E374 (E374D) or E399 (E399D), the channel remained sensitive to Mg^{2+} (Fig. 4a), indicating that Mg^{2+} could still bind to the site. However, Mg^{2+} sensitivity was reduced by both mutations owing to the change in size of the side chain. The effects of E374D and E399D are not equal (Fig. 4a). This is possibly due to the coordination of the main chain carbonyl of Q397 pulling the metal closer to the backbone of β C (Fig. 3) and thus helping to better accommodate the space change caused by E399D. Notably, when E374 or E399 is replaced by non-charged (Ala or Gln) or positively charged residues (Arg), an increase of $[Mg^{2+}]_i$ from 0 to 10 mM at 0 $[Ca^{2+}]_i$ still activated the channel to a small but observable extent (Figs 2c and 4a). As these mutations should destroy the low-affinity metal-binding site, such a residual sensitivity to Mg^{2+} is probably derived from Mg^{2+} binding to the high-affinity Ca^{2+} site, which subsequently activated the channel. Consistent with this idea, at 110 μ M $[Ca^{2+}]_i$, which is saturating for the high-affinity Ca^{2+} site⁸, Mg^{2+} could no longer activate these mutant channels (Figs 2c and 4a). This idea is further supported by the result that the mSlo3 channel, which also lacks Ca^{2+} sensitivity, was not activated by Mg^{2+} at 0 $[Ca^{2+}]_i$ (ref. 9) (Fig. 1c).

Figure 4b shows that of the mutations that impaired or abolished Mg^{2+} sensitivity, none affected the free energy provided by Ca^{2+} binding towards the activation of the channel when $[Ca^{2+}]_i$ increased from 0 to 110 μ M, at an intracellular Mg^{2+} concentration

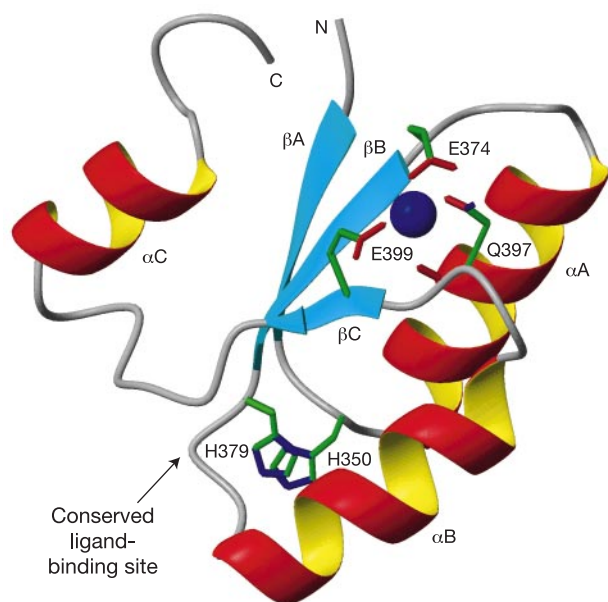


Figure 3 Structure of the Mg^{2+} -binding site. The blue sphere indicates a bound metal ion. Residues are labelled with their position in mSlo1. The conserved ligand-binding site in most other ligand-binding proteins with a Rossman fold is labelled with an arrow.

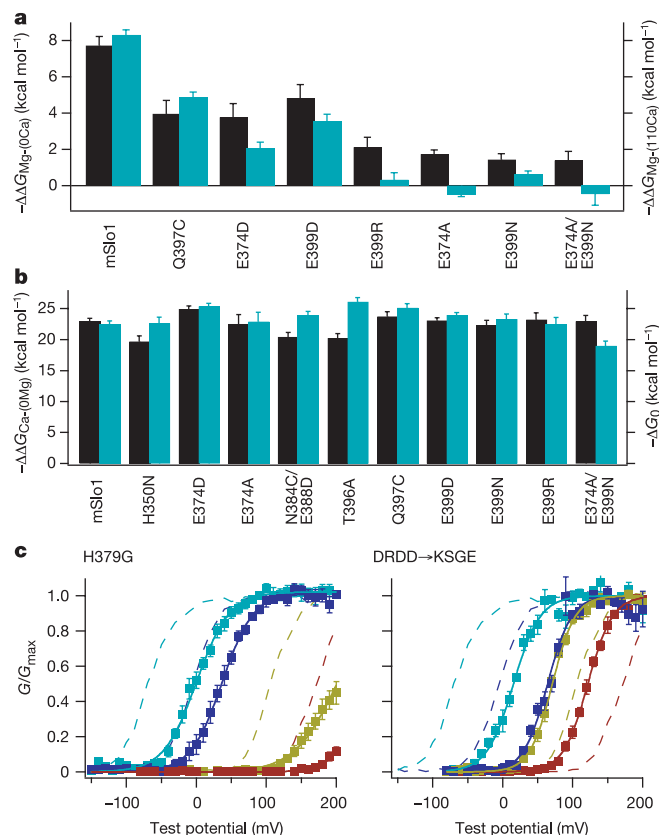


Figure 4 Results of site-directed mutations that affect channel gating. **a**, Mutations on Mg^{2+} -coordinating residues. The graph shows the free energy provided by Mg^{2+} binding towards the activation of channels when $[Mg^{2+}]_i$ increases from 0 to 10 mM, measured at 0 (left y axis, black columns) and 110 μ M (right y axis, green columns) $[Ca^{2+}]_i$ ($\Delta\Delta G_{Mg-Ca}$). **b**, Mutations affecting Mg^{2+} sensitivity. The graph shows the free energy increase in the absence of Ca^{2+} and Mg^{2+} at 0 mV (ΔG_0 , green columns), and the free energy provided by Ca^{2+} binding towards the activation of channels when $[Ca^{2+}]_i$ increases from 0 to 110 μ M at 0 $[Mg^{2+}]_i$ ($\Delta\Delta G_{Ca-0Mg}$, black columns). **c**, Mean G - V relations of H379G and DRDD/KSGE channels. $[Mg^{2+}]_i$ and $[Ca^{2+}]_i$ are colour-coded as in Fig. 1c. Dashed lines are G - V relations of mSlo1 for comparison. $n = 4-7$ patches for each mutation.

of 0 ($\Delta\Delta G_{\text{Ca}-(\text{OMg})}$; see also Methods). The measurement of $\Delta\Delta G_{\text{Ca}-(\text{OMg})}$ for mutation H379G, which also reduced Mg^{2+} sensitivity (Fig. 2b), is not shown on Fig. 4b because the G - V relation was shifted to right and could not be measured accurately (Fig. 4c). Instead, a value of $\Delta\Delta G_{\text{Ca}-(\text{OMg})} = 12.5 \pm 0.7 \text{ kcal mol}^{-1}$ ($n = 6$) was measured between 2 and $110 \mu\text{M}$ $[\text{Ca}^{2+}]_i$ and compared with that of mSlo1 channels ($12.0 \pm 0.4 \text{ kcal mol}^{-1}$, $n = 3$); this treatment showed no significant difference. Therefore, the structural components that are affected by these mutations are involved in Mg^{2+} - but not Ca^{2+} -dependent activation. On the other hand, when residues in the interloop connecting αA - βB of mSlo1 were mutated to corresponding mSlo3 residues (DRDD $\rightarrow$ KSGE, Figs 2a and 3), the free energy provided by Ca^{2+} -binding towards the activation of the channel was reduced to about one-third of that for mSlo1 (Fig. 4c). This result indicates an important role of this loop in Ca^{2+} -dependent activation. However, although this loop is spatially close to the Mg^{2+} -binding site, mutation DRDD $\rightarrow$ KSGE had no effect on Mg^{2+} sensitivity (Fig. 2b). Similarly, when the tail domain of mSlo1 is substituted by the tail of mSlo3, the channel loses Ca^{2+} sensitivity but retains an intact Mg^{2+} sensitivity^{9,18}. These results indicate that the tail domain and the interloop connecting αA - βB in the RCK domain may be involved in Ca^{2+} - but not Mg^{2+} -dependent activation. Thus, distinct structural components are responsible for Ca^{2+} - and Mg^{2+} -dependent activation of BK channels. Not only are the binding sites different, but also the energy provided by metal binding affects different sets of local conformational changes that eventually lead to channel opening.

The results from chimera studies (Fig. 1) demonstrate that only a small fragment of mSlo1 is necessary and sufficient for restoring Mg^{2+} sensitivity in mSlo3 channels despite numerous other amino acid sequence differences between the two. Within this fragment, H350, H379 and N384 are located in the conserved active site of other proteins that adopt a Rossman fold^{12,25} (arrow in Fig. 3). Mutations of these residues affect Mg^{2+} sensitivity (Fig. 2b). These results mark the conserved active site as the only area away from the Mg^{2+} binding site that is important for Mg^{2+} sensitivity. These residues do not act independently to affect Mg^{2+} sensitivity, because the reduction of free energy provided by Mg^{2+} binding towards the activation of the channel caused by each individual mutation does not add up to the reduction caused by chimera C31-I, which includes all of these individual mutations (Figs 1c and 2b)²⁸. Consistent with these results, structural analysis indicates that the side chains of H350 and H379 are close to each other, possibly interacting by stacking of their aromatic rings (Fig. 3). Thus, it appears that in BK channels Mg^{2+} binding at the N terminus of β -strands places an energetic constraint on the conformational change at the conserved active site located at the C terminus of β -strands, thereby activating the channel. Given the close proximity between the RCK domain and the S6 helix that forms the activation gate¹³⁻¹⁵ (Fig. 1a), it is possible that the conserved active site directly interacts with the activation gate. □

Methods

Mutagenesis and expression

All channel constructs were made from the mbr5 clone of mSlo1 (ref. 16) and the complementary DNA of mSlo3 (ref. 20). In C31-I, amino acids 348–386 of mSlo1 are substituted by amino acids 337–375 of mSlo3. In C31-II, amino acids 385–420 of mSlo1 are substituted by amino acids 374–409 of mSlo3. In C13, amino acids 361–405 of mSlo3 are substituted by amino acids 372–416 of mSlo1. The polymerase chain reaction (PCR)-amplified regions of all mutants were verified by sequencing. A PCR error in chimera C31-I resulted in an additional mutation V553G. As V553 is not conserved between the Mg^{2+} -sensitive mSlo1 and dSlo, the additional mutation was not corrected. RNA was transcribed *in vitro* with T3 polymerase (Ambion). We injected 0.05–50 ng of RNA into each *Xenopus laevis* oocyte 2–6 days before recording.

Electrophysiology

Macroscopic currents were recorded from inside-out patches formed with borosilicate pipettes of 0.9–1.8 M Ω resistance. Data were acquired using an Axopatch 200-B patch clamp amplifier (Axon Instruments) and pulse acquisition software (HEKA Elektronik).

Records were digitized at 20- μs intervals and low-pass-filtered at 10 KHz with the 4 pole Bessel filter of Axopatch. The pipette solution contained (in mM): 140 potassium methanesulphonic acid, 20 HEPES, 2 KCl, 2 MgCl₂, pH 7.20. The basal internal solution contained (in mM): 140 potassium methanesulphonic acid, 20 HEPES, 2 KCl, 1 EGTA, pH 7.20. The '0 $[\text{Ca}^{2+}]_i$ ' solution was the same as the basal internal solution except that it contained 5 mM EGTA, having a free $[\text{Ca}^{2+}]_i$ of approximately 0.5 nM that was too low to affect mSlo1 channel activation⁸. CaCl_2 and MgCl_2 were added to internal solutions to give the appropriate free $[\text{Ca}^{2+}]_i$ and $[\text{Mg}^{2+}]_i$ (ref. 9). We obtained all recordings at room temperature (22–24 °C).

Analysis

Relative conductance was determined by measuring tail current amplitudes at -50 mV . G - V relations were fitted with the Boltzmann distribution $G/G_{\text{max}} = 1/[1 + \exp(\Delta G_{\text{Act}}/kT)]$, where k is Boltzmann's constant, T is absolute temperature, and ΔG_{Act} is the free energy change of channel opening. Because voltage, Ca^{2+} and Mg^{2+} open the activation gate independently, ΔG_{Act} is the sum of energy increase provided by voltage ($\Delta G_V = -zeV$, where e is the elementary charge and z is the number of equivalent charges), Ca^{2+} and Mg^{2+} binding (ΔG_{Ca} , ΔG_{Mg}), and that in the absence of Ca^{2+} and Mg^{2+} at 0 mV (ΔG_0)^{9,11}; so $\Delta G_{\text{Act}} = \Delta G_V + \Delta G_{\text{Ca}} + \Delta G_{\text{Mg}} + \Delta G_0$.

The change in Ca^{2+} - or Mg^{2+} -binding contribution to ΔG_{Act} as a result of an increase of $[\text{Ca}^{2+}]_i$ or $[\text{Mg}^{2+}]_i$ ($\Delta\Delta G_{\text{Ca}}$ or $\Delta\Delta G_{\text{Mg}}$) was then calculated based on the shift of the G - V relation: $\Delta\Delta G_{\text{Ca}} = -\Delta(zeV_{1/2})$ or $\Delta\Delta G_{\text{Mg}} = -\Delta(zeV_{1/2})$, where $V_{1/2}$ is the voltage at half maximum of the G - V relation. Error bars in all figures show standard error of means.

Molecular modelling

The structure of the RCK domain was generated by residual replacement based on the coordinates of the crystal structure of the RCK domain from the *E. coli* K⁺ channel¹². Magnesium ion was manually docked onto the potential Mg^{2+} -binding site as derived from mutation results. Then, the energy minimization was performed by fixing all backbone atoms and using the Discover Package with Insight II interface (Molecular Simulation). Structural diagrams were made using MolMol²⁹.

Received 25 April; accepted 17 June 2002; doi:10.1038/nature00941.

- Brenner, R. et al. Vasoregulation by the $\beta 1$ subunit of the calcium-activated potassium channel. *Nature* **407**, 870–876 (2000).
- Robitaille, R., Garcia, M. L., Kaczorowski, G. J. & Charlton, M. P. Functional colocalization of calcium and calcium-gated potassium channels in control of transmitter release. *Neuron* **11**, 645–655 (1993).
- Isaacson, J. S. & Murphy, G. J. Glutamate-mediated extrasynaptic inhibition: direct coupling of NMDA receptors to $\text{Ca}(2+)$ -activated K⁺ channels. *Neuron* **31**, 1027–1034 (2001).
- Fettiplace, R. & Fuchs, P. A. Mechanisms of hair cell tuning. *Annu. Rev. Physiol.* **61**, 809–834 (1999).
- Rosenblatt, K. P., Sun, Z. P., Heller, S. & Hudspeth, A. J. Distribution of Ca^{2+} -activated K⁺ channel isoforms along the tonotopic gradient of the chicken's cochlea. *Neuron* **19**, 1061–1075 (1997).
- Pallotta, B. S., Magleby, K. L. & Barrett, J. N. Single channel recordings of Ca^{2+} -activated K⁺ currents in rat muscle cell culture. *Nature* **293**, 471–474 (1981).
- Marty, A. Ca-dependent K channels with large unitary conductance in chromaffin cell membranes. *Nature* **291**, 497–500 (1981).
- Cui, J., Cox, D. H. & Aldrich, R. W. Intrinsic voltage dependence and Ca^{2+} regulation of mSlo large conductance Ca-activated K⁺ channels. *J. Gen. Physiol.* **109**, 647–673 (1997).
- Shi, J. & Cui, J. Intracellular $\text{Mg}(2+)$ enhances the function of BK-type $\text{Ca}(2+)$ -activated K(+) channels. *J. Gen. Physiol.* **118**, 589–606 (2001).
- Zhang, X., Solaro, C. R. & Lingle, C. J. Allosteric regulation of BK channel gating by $\text{Ca}(2+)$ and $\text{Mg}(2+)$ through a nonselective, low affinity divalent cation site. *J. Gen. Physiol.* **118**, 607–636 (2001).
- Cui, J. & Aldrich, R. W. Allosteric linkage between voltage and $\text{Ca}(2+)$ -dependent activation of BK-type mSlo1 K(+) channels. *Biochemistry* **39**, 15612–15619 (2000).
- Jiang, Y., Pico, A., Cadene, M., Chait, B. T. & MacKinnon, R. Structure of the RCK domain from the *E. coli* K⁺ channel and demonstration of its presence in the human BK channel. *Neuron* **29**, 593–601 (2001).
- Yellen, G. The moving parts of voltage-gated ion channels. *Q. Rev. Biophys.* **31**, 239–295 (1998).
- Perozo, E., Cortes, D. M. & Cuello, L. G. Structural rearrangements underlying K⁺-channel activation gating. *Science* **285**, 73–78 (1999).
- Johnson, J. P. Jr & Zagotta, W. N. Rotational movement during cyclic nucleotide-gated channel opening. *Nature* **412**, 917–921 (2001).
- Butler, A., Tsunoda, S., McCobb, D. P., Wei, A. & Salkoff, L. mSlo, a complex mouse gene encoding 'maxi' calcium-activated potassium channels. *Science* **261**, 221–224 (1993).
- Wei, A., Solaro, C., Lingle, C. & Salkoff, L. Calcium sensitivity of BK-type KCa channels determined by a separable domain. *Neuron* **13**, 671–681 (1994).
- Schreiber, M., Yuan, A. & Salkoff, L. Transplantable sites confer calcium sensitivity to BK channels. *Nature Neurosci.* **2**, 416–421 (1999).
- Bian, S., Favre, I. & Moczydlowski, E. Ca^{2+} -binding activity of a COOH-terminal fragment of the *Drosophila* BK channel involved in Ca^{2+} -dependent activation. *Proc. Natl Acad. Sci. USA* **98**, 4776–4781 (2001).
- Schreiber, M. et al. Slo3, a novel pH-sensitive K⁺ channel from mammalian spermatocytes. *J. Biol. Chem.* **273**, 3509–3516 (1998).
- Adelman, J. P. et al. Calcium-activated potassium channels expressed from cloned complementary DNAs. *Neuron* **9**, 209–216 (1992).
- Milkman, R. An *Escherichia coli* homologue of eukaryotic potassium channel proteins. *Proc. Natl Acad. Sci. USA* **91**, 3510–3514 (1994).
- Stock, J. B., Surette, M. G., McCleary, W. R. & Stock, A. M. Signal transduction in bacterial chemotaxis. *J. Biol. Chem.* **267**, 19753–19756 (1992).
- Lee, J. O., Rieu, P., Arnaout, M. A. & Liddington, R. Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* **80**, 631–638 (1995).
- Brandeen, C. I. Relation between structure and function of alpha/beta-proteins. *Q. Rev. Biophys.* **13**, 317–338 (1980).

26. Needham, J. V., Chen, T. Y. & Falke, J. J. Novel ion specificity of a carboxylate cluster Mg(II) binding site: strong charge selectivity and weak size selectivity. *Biochemistry* **32**, 3363–3367 (1993).
27. Flatman, P. W. Mechanisms of magnesium transport. *Annu. Rev. Physiol.* **53**, 259–271 (1991).
28. Horowitz, A. & Fersht, A. R. Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins. *J. Mol. Biol.* **214**, 613–617 (1990).
29. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55 (1996).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

The mSlo1 and mSlo3 clones were provided by L. Salkoff. We thank S. Chen, S. W. Jones and R. Aldrich for comments on the manuscript. This work was supported by grants from the NIH (to J.Q. and J.C.), the American Heart Association and the Whitaker Foundation (to J.C.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C. (e-mail: jxc93@cwru.edu).

Multiple regulatory sites in large-conductance calcium-activated potassium channels

Xiao-Ming Xia, Xuhui Zeng & Christopher J. Lingle

Department of Anesthesiology, Washington University School of Medicine, Box 8054, St. Louis, Missouri 63110, USA

Large conductance, Ca^{2+} - and voltage-activated K^+ channels (BK) respond to two distinct physiological signals—membrane voltage and cytosolic Ca^{2+} (refs 1, 2). Channel opening is regulated by changes in Ca^{2+} concentration spanning 0.5 μM to 50 mM (refs 2–5), a range of Ca^{2+} sensitivity unusual among Ca^{2+} -regulated proteins. Although voltage regulation arises from mechanisms shared with other voltage-gated channels^{6–8}, the mechanisms of Ca^{2+} regulation remain largely unknown. One potential Ca^{2+} -regulatory site, termed the 'Ca²⁺ bowl', has been located to the large cytosolic carboxy terminus^{9–11}. Here we show that a second region of the C terminus, the RCK domain (regulator of conductance for K^+ (ref. 12)), contains residues that define two additional regulatory effects of divalent cations. One site, together with the Ca^{2+} bowl, accounts for all physiological regulation of BK channels by Ca^{2+} ; the other site contributes to effects of millimolar divalent cations that may mediate physiological regulation by cytosolic Mg^{2+} (refs 5, 13). Independent regulation by multiple sites explains the large concentration range over which BK channels are regulated by Ca^{2+} . This allows BK channels to serve a variety of physiological roles contingent on the Ca^{2+} concentration to which the channels are exposed^{14,15}.

A recent advance in understanding the regulation of BK channels by Ca^{2+} was the demonstration that distinct, independent binding sites for Ca^{2+} may explain the range of intracellular Ca^{2+} concentration ($[\text{Ca}^{2+}]_i$) that regulates channel opening^{5,13}. Over the physiological range of $[\text{Ca}^{2+}]_i$, relatively Ca^{2+} -selective site(s) mediate channel regulation. In contrast, the effects of $[\text{Ca}^{2+}]_i$ in the millimolar range reflect a lower sensitivity site showing selectivity for Ca^{2+} and Mg^{2+} . Here we define principal structural elements required for each Ca^{2+} -dependent regulatory effect. BK channel α -subunits are encoded by a single *Slo1* gene^{16,17}, with a

functional channel arising from the tetrameric assembly of four α -subunits¹⁸. Although the transmembrane segments (S1–S6, Fig. 1a) of each α -subunit^{17,19} share homology with voltage-gated K^+ channels, each α -subunit uniquely contains an extensive C terminus with four additional hydrophobic segments (S7–S10). The C terminus is probably composed of two modular units, as expression in oocytes of separate messages for S0–S8 and S9–S10 peptides produces channels identical to wild-type BK channels²⁰. Whereas the Ca^{2+} bowl is contained within the S9–S10 peptide, the C-terminal structure with the S7–S8 segments contains the RCK domain, a second potential regulatory element (Fig. 1a, b) that exhibits extensive homology with a number of bacterial K^+ channels. The structures of two prokaryotic RCK domains have been determined^{12,21} and RCK domains may contain binding sites for a variety of regulatory ligands, including nucleotides²² and cations²¹.

To address the role of the RCK domain in BK channel regulation, we took advantage of the fact that a homologue of *Slo1*, the *Slo3* pH-sensitive K^+ channel, lacks Ca^{2+} -dependent regulation^{13,23,24}. We focused on residues near folds in the RCK domain that contribute to nucleotide binding in prokaryotic homologues (Fig. 1c). We were

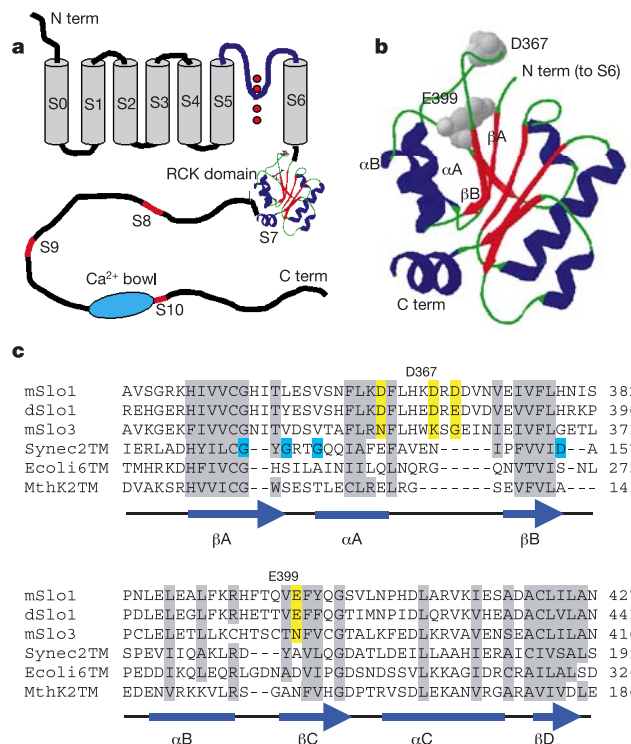


Figure 1 Schematic representation of the Slo1 α -subunit. **a**, The Slo1 α -subunit shares a common transmembrane topology (S1–S6) with voltage-dependent K^+ channels, and contains a unique amino-terminal S0 segment and an extensive cytosolic C-terminal elaboration with hydrophobic segments S8–S10 (red), the Ca^{2+} bowl, and the RCK domain. Segment S7 is contained in the RCK structure¹². **b**, Structural approximation of the BK channel RCK domain based on the RCK domain of the *E. coli* K^+ channel¹². The position of the linker between helix α A and β B is uncertain. **c**, Sequence alignment within the RCK domain of Slo1 and Slo3 subunits and three bacterial K^+ channels¹². Grey indicates semi-conserved residues; blue shows the NAD-binding motif present in some bacterial homologues; and yellow indicates negative residues in Slo1 that are non-conserved in Slo3. mSlo1, mouse BK channel, *Mus musculus* (GenBank accession number: 6754435); dslo1, *Drosophila* BK channel, *D. melanogaster* (GenBank accession number: 7301192); mSlo3, mouse Slo3 channel, *M. musculus* (GenBank accession number: 6680542); Synec2TM, *Synectocystis* sp. (GenBank accession number: 7447543); Ecoli6TM, *E. coli* (GenBank accession number: 400124); MthK2TM, *M. thermotrophicum* (GenBank accession number: 2622639).

particularly interested in negatively charged residues conserved in Ca^{2+} -dependent forms of *Slo1* (for example, *Drosophila* and mouse *Slo1* (*dSlo1* and *mSlo1*, respectively)), but not in *Slo3*. To assess whether a construct exhibits altered Ca^{2+} -dependent gating, we examined the activation characteristics over a range of $[\text{Ca}^{2+}]_i$ and $[\text{Mg}^{2+}]_i$. In accordance with allosteric models of BK channel activation^{4,5,7,13}, this approach reveals to what extent a structural alteration influences divalent cation binding, movement of the voltage sensors, or the closed–open conformational equilibria.

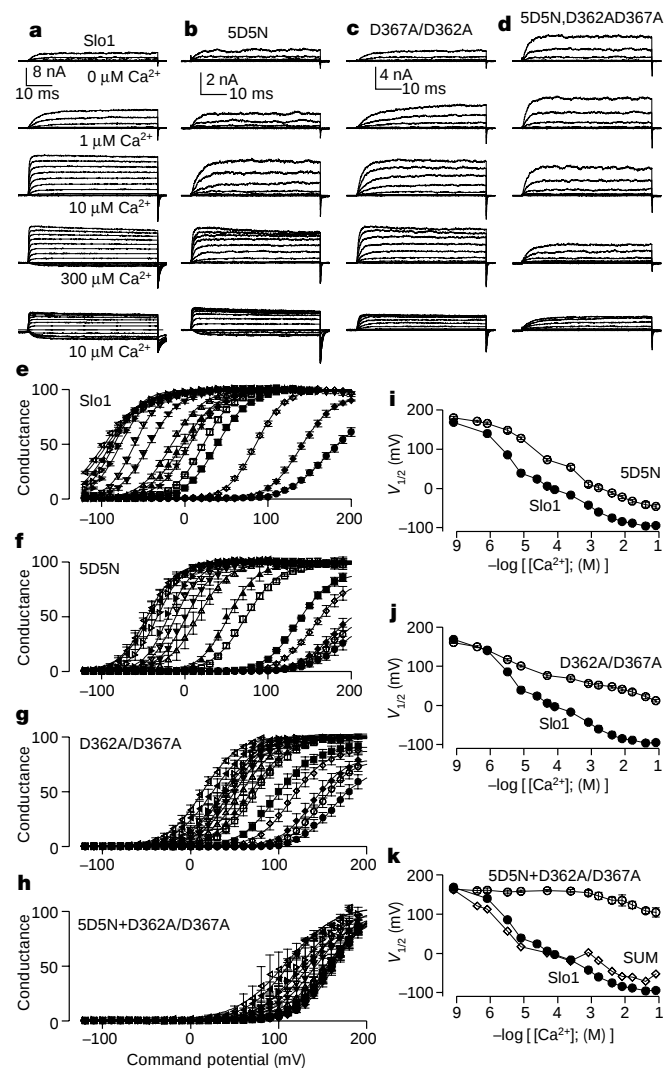


Figure 2 The Ca^{2+} bowl and D362A/D367A account for all physiologically relevant Ca^{2+} -dependent regulation of BK channels. **a–d**, Traces show currents from inside-out patches for *Slo1*, 5D5N, D362A/D367A and the combined mutation 5D5N + D362A/D367A with cytosolic $[\text{Ca}^{2+}]_i$, as shown. Voltage steps were from -100 through to $+160$ in 20 -mV increments following 40 ms at -180 mV, with tail currents at -120 mV. **e–h**, G - V curves were generated over $[\text{Ca}^{2+}]_i$, from 0 to 100 mM for *Slo1* and the indicated mutants. Solid lines are fits of equation (1) (see Methods). $[\text{Ca}^{2+}]_i$ are: filled circles, 0 μM ; open circles, 0.5 μM ; filled diamonds, 1 μM ; open diamonds, 4 μM ; filled squares, 10 μM ; open squares, 60 μM ; stars, 100 μM ; filled triangles, 300 μM ; open triangles, 1 mM; filled triangles (inverted), 2 mM; open triangles (inverted), 5 mM; filled triangles (facing right), 10 mM; open triangles (facing right), 20 mM; filled triangles (facing left), 50 mM; open triangles (facing left), 100 mM. **i**, Activation $V_{1/2}$ versus $[\text{Ca}^{2+}]_i$ for wild-type BK currents and for mutation 5D5N. **j**, **k**, Wild-type (*Slo1*) versus D362A/D367A (**j**) and wild-type versus combined mutation of the Ca^{2+} bowl and D362A/D367A (**k**, 5D5N + D362A/D367A). The $V_{1/2}$ shift predicted for the case in which the Ca^{2+} bowl and D362A/D367A each independently regulate gating is also shown (SUM).

We first compared the Ca^{2+} sensitivity of wild-type BK currents arising from *Slo1* α -subunits to that arising from *Slo1* subunits with a mutated Ca^{2+} bowl region (5D5N; see Methods). This mutation abolishes function of the Ca^{2+} bowl^{9,11}. Currents resulting from the 5D5N construct still exhibit substantial regulation by Ca^{2+} , although the ability of $[\text{Ca}^{2+}]_i$ at less than 10 μM to cause channel activation is reduced relative to wild-type currents (Fig. 2a, b). As seen in a comparison of conductance–voltage (G - V) curves (Fig. 2e, f), both constructs exhibit similar activation in the absence of Ca^{2+} , but $[\text{Ca}^{2+}]_i$ below 10 μM has less effect in the 5D5N construct.

We next examined whether the RCK domain contributes to Ca^{2+} -dependent regulation of current that persists after mutation of the Ca^{2+} bowl. Of the prokaryotic channels containing an RCK domain, only some contain the conserved motif between β -sheet A and α -helix B (Fig. 1c) that defines nucleotide binding²². This region of *Slo1* and *Slo3* exhibits no obvious residues that might account for a difference in Ca^{2+} sensitivity. However, at the end of α A and through the linker up to β B, there are three aspartate residues (D362, D367 and D369 in *mSlo1*) conserved in *Slo1* forms, but which are neutral or positively charged in *Slo3*. Compared with bacterial homologues¹², the linker between α A and β B in the *Slo* family of subunits contains five extra residues (Fig. 1c), suggesting an important functional role for this segment.

The point mutations D362A and D369A each resulted in currents with Ca^{2+} sensitivity similar to wild-type currents, with a slight effect of D362A at lower $[\text{Ca}^{2+}]_i$. In contrast, D367A produced a marked reduction in the ability of Ca^{2+} to shift gating. Constructs D362A/D367A and D362A/D367A/D369A exhibited behaviour similar to D367A alone. We focused on the D362A/D367A double mutant. Currents arising from D362A/D367A exhibited reduced Ca^{2+} -dependent regulation of activation (Fig. 2c). The resulting G - V curves (Fig. 2g) show that the full shift in gating over all $[\text{Ca}^{2+}]_i$ with the D362A/D367A mutation is less than that in 5D5N, although D362A/D367A exhibits more sensitivity to $[\text{Ca}^{2+}]_i$ below 10 μM (Fig. 2f, g). When both the Ca^{2+} bowl and RCK

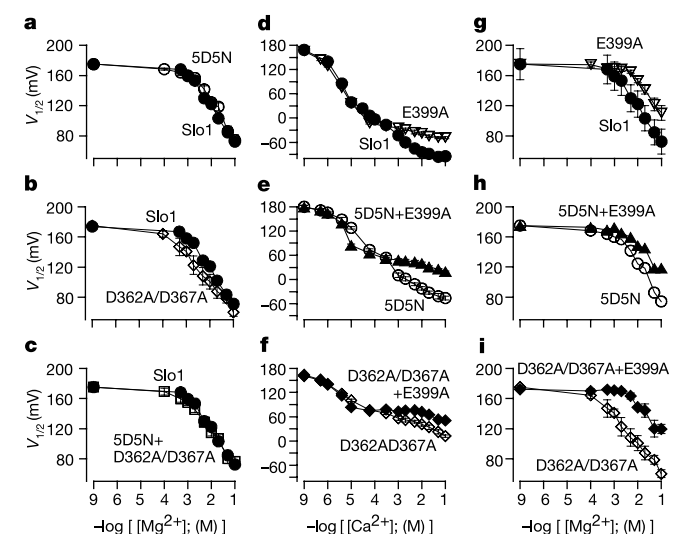


Figure 3 The RCK domain also mediates effects of high concentrations of Ca^{2+} and Mg^{2+} . **a–c**, $V_{1/2}$ as a function of $[\text{Mg}^{2+}]_i$ for pairs of site mutations. **d–f**, $V_{1/2}$ as a function of $[\text{Ca}^{2+}]_i$ for different mutations with and without E399A. **g–i**, $V_{1/2}$ as a function of $[\text{Mg}^{2+}]_i$ for different mutations with and without E399A. *Slo1* is indicated by filled circles. $V_{1/2}$ shifts resulting from changing $[\text{Ca}^{2+}]_i$ from 1 to 100 mM: *Slo1*, 51.5 ± 2.4 mV; plus E399A*, 22.3 ± 1.5 mV; 5D5N, 62.2 ± 4.9 mV, plus E399A*, 26.0 ± 3.5 mV; D362A/D367A, 48.0 ± 3.9 mV, plus E399A, 23.5 ± 5.5 mV. $V_{1/2}$ shifts for 0 – 100 mM $[\text{Mg}^{2+}]_i$: *Slo1*, 109.8 ± 2.8 mV, plus E399A*, 63.9 ± 6.6 mV; 5D5N, 108.9 ± 5.6 mV, plus E399A*, 56.7 ± 4.3 ; D362A/D367A, 112.0 ± 4.2 mV, plus E399A*, 75.5 ± 6.3 mV. Asterisk, $P < 0.001$; double asterisk, $P < 0.002$.

domain residues were mutated simultaneously (5D5N plus D362A/D367A), the channels were still gated by voltage in a fashion similar to wild-type currents, but Ca^{2+} had essentially no effect on gating at concentrations up to 1 mM (Fig. 2d, h). Together, the loci defined by 5D5N and D362A/D367A account for all Ca^{2+} -dependent regulation of BK channels over the physiologically important range of $[\text{Ca}^{2+}]_i$.

For BK currents, the voltage of half activation ($V_{1/2}$) of conductance at different $[\text{Ca}^{2+}]_i$ provides a useful indicator of the apparent Ca^{2+} -dependence of activation. For all four constructs, the $V_{1/2}$ for activation at 0 Ca^{2+} remained within 20 mV of each other (Fig. 2i–k), indicative that these mutations were not altering strictly voltage-dependent equilibria. Allosteric models predict that, if two

loci reflect the independent contribution of two Ca^{2+} -regulatory elements, the magnitude of the shifts in $V_{1/2}$ resulting from the effects of each individual site should sum to be similar to the effect of the combined mutation^{5,7,13} (equation (3); see also Methods). Therefore, we summed the residual shift in $V_{1/2}$ after mutation of both sites with the increments in $V_{1/2}$ attributable to the Ca^{2+} bowl and to the D362/D367 site (Fig. 2k, open diamonds). The correspondence between wild-type values of $V_{1/2}$ and those predicted from adding the 5D5N and D362A/D367A mutations suggests that each locus acts in a relatively independent way to regulate BK gating^{4,5,13}.

With $[\text{Ca}^{2+}]_i$ above 1 mM, the 5D5N plus D362A/D367A mutants show an additional 50–60-mV negative shift in $V_{1/2}$, despite the lack of effect of $[\text{Ca}^{2+}]_i$ below 1 mM (Fig. 2k). This effect is similar to previously described low-affinity effects of Mg^{2+} (refs 5, 13, 25, 26) and Ca^{2+} (refs 5, 13) on BK activation. We therefore examined the ability of Mg^{2+} to produce shifts in activation in 5D5N and D362A/D367A, and the combined 5D5N plus D362A/D367A mutations. In all cases (Fig. 3a–c), the Mg^{2+} -induced shift in activation was similar to that observed in wild-type currents, indicative that the mechanism responsible for the Mg^{2+} effect was intact. This also indicates that 5D5N and D362A/D367A do not nonspecifically alter the ability of the channel to be regulated.

We next examined the role of the RCK domain on effects of millimolar concentrations of divalent cations. E399 is conserved in mammalian, fly and worm *Slo1* sequences, but not in *Slo3*. Its location on the β C sheet (Fig. 1b–c) may position it near the putative ligand-binding region. The point mutation E399A resulted in currents with reduced sensitivity only above 1 mM Ca^{2+} (Fig. 3d). Furthermore, E399A reduced the ability of Mg^{2+} to shift gating (Fig. 3g). E399A, when paired either with 5D5N (Fig. 3e, h) or D362A/D367A (Fig. 3f, i), reduced the effects of both millimolar Ca^{2+} and millimolar Mg^{2+} to shift gating. In all constructs containing E399A, the $V_{1/2}$ shift resulting from increasing $[\text{Ca}^{2+}]_i$ from 1 to 100 mM is reduced by about 30–40 mV. The $V_{1/2}$ shift produced by raising $[\text{Mg}^{2+}]_i$ from 0 to 100 mM was reduced from about 110 mV to about 60 mV. The similarity of the effect of E399A in all constructs suggests that the divalent cation effect influenced by E399 acts independently of other Ca^{2+} -regulatory elements. Thus, E399 seems to define an independent lower-affinity, divalent cation-regulatory effect.

If E399 accounts for most of the additional effects of Ca^{2+} at millimolar concentrations, the triple mutant construct 5D5N plus D362A/D367A, plus E399A might be expected to abolish all Ca^{2+} sensitivity. Remarkably, currents resulting from the triple mutant did not simply exhibit insensitivity to Ca^{2+} , but were markedly suppressed as Ca^{2+} was elevated from 300 μM through to 100 mM (Fig. 4a). Although rapid block of open channels by millimolar concentrations of Ca^{2+} or Mg^{2+} is a common feature of these channels^{5,13,27}, such open channel block is rapidly relieved on hyperpolarization, with little steady-state block at potentials negative to –100 mV. Yet, for the triple mutant construct, currents are strongly inhibited in a voltage-independent fashion. Furthermore, the Ca^{2+} -dependent inhibition is associated with a rightward shift in the G – V curve compared with that in 0 Ca^{2+} (Fig. 4c). In contrast, in the triple mutation constructs Mg^{2+} still produces shifts in activation to more negative potentials (Fig. 4b).

What is the mechanism by which E399A in conjunction with mutation of the other two loci produces this suppression of current? The energetic additivity of the effects of 5D5N, D362A/D367A (Fig. 2k) and E399A (Fig. 3d–f) indicates that none of the mutations produce a nonspecific suppression in the ability of current to be regulated by Ca^{2+} . An alternative explanation is that one mutation has altered the properties of a Ca^{2+} -binding site, such that Ca^{2+} now binds more tightly to the closed channel than to the open channel. For an allosteric model in which distinct regulatory sites independently regulate the channel closed-to-open equilibria^{5,7,13},

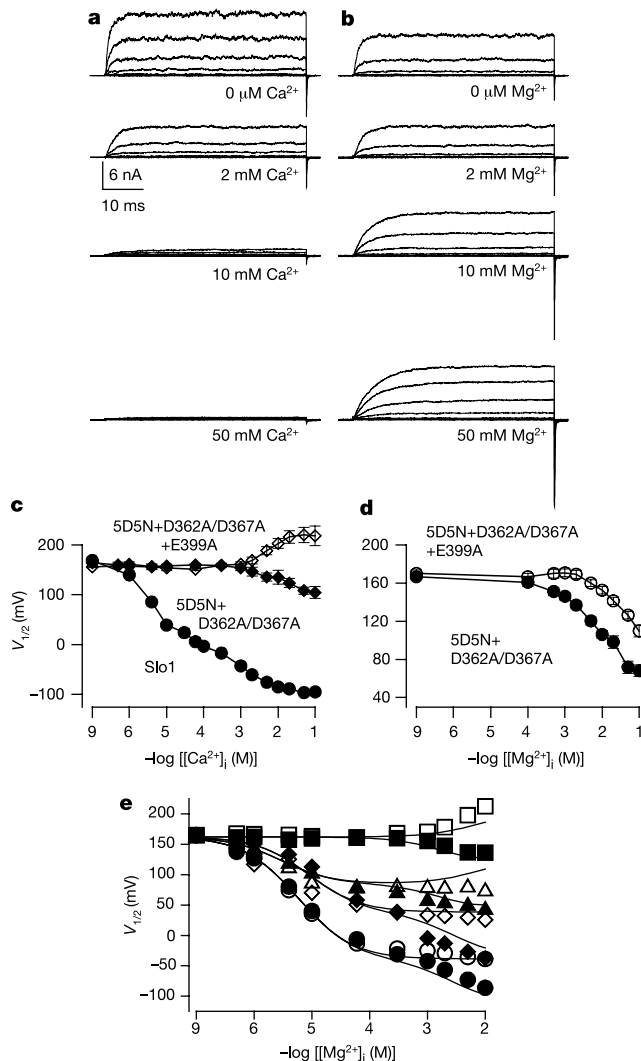


Figure 4 Mutation of all three regulatory elements produces Ca^{2+} -dependent suppression of BK current. **a** Currents resulting from the triple mutation 5D5N + D362A/D367A + E399A were activated as in Fig. 1. **b**, Mg^{2+} shifts activation of currents resulting from the triple mutation. **c**, $V_{1/2}$ is plotted versus $[\text{Ca}^{2+}]_i$ for wild type (Slo1), 5D5N + D362A/D367A and 5D5N + D362A/D367A + E399A. **d**, $V_{1/2}$ is plotted versus $[\text{Mg}^{2+}]_i$ for 5D5N + D362A/D367A, and 5D5N + D362A/D367A + E399A. **e**, An allosteric model with three independent Ca^{2+} -binding sites (equation (3)) was used to fit simultaneously the relationship between $V_{1/2}$ and $[\text{Ca}^{2+}]_i$ for all constructs. Parameter values are given in the Methods. Slo1, filled circles; E399A, open circles; 5D5N, filled diamonds; 5D5N + E399A, open diamonds; D362A/D367A, filled triangles; D362A/D367A + E399A, open triangles; 5D5N + D362A/D367A, filled squares; 5D5N + D362A/D367A + E399A, open squares.

such a change in affinity will suppress current activation and shift G - V curves to the right. In accordance with general pharmacological principles, in the triple mutant Ca^{2+} now appears to act as an inverse agonist at one of the Ca^{2+} -binding sites. We propose that D362/D367 is probably responsible for this effect, although this will require more investigation. We have evaluated the behaviour of the wild-type currents and the various mutants in terms of a three-regulatory-site model (Fig. 4e). This model adequately accounts for many of the principal features of the behaviour of $V_{1/2}$ for all of the constructs, and supports the validity of the allosteric model of regulation of BK gating^{4,5,7,13,28}.

These results show that the RCK domain is critical to Ca^{2+} -dependent regulation of BK channels and participates in both higher (D362 and D367) and lower (E399) affinity effects of divalent cations. Irrespective of whether any of the critical residues in the RCK domain mediates binding with divalent cations, mutation of these residues clearly disrupts channel regulatory mechanisms that are intimately associated with specific aspects of Ca^{2+} -dependent activation. Do these loci actually reflect divalent cation binding sites? We currently consider this an attractive hypothesis. First, the principal residues are situated in positions with proximity to the nucleotide-binding sites described in other homologues, suggesting a similar regulatory role. Second, the principal residues are anionic. Third, the results suggest that a specific mutation has resulted in a change in Ca^{2+} -binding affinities, such that Ca^{2+} now acts to inhibit current activation. If a mutation simply disrupted transduction of a process regulated by a Ca^{2+} -binding site located elsewhere, a change in molecular binding affinities would not be expected. However, it remains plausible that the mutations studied here have simply disrupted the ability of the RCK domain to exert regulatory effects mediated by Ca^{2+} binding elsewhere.

The presence of multiple Ca^{2+} -dependent regulatory domains on the Slo1 α -subunit, each with different affinities for Ca^{2+} , provides a mechanism by which the broad concentration range of Ca^{2+} -dependence of BK currents can be explained. Although further analysis may reveal some linkage between the allosteric effects mediated by each regulatory element, it is remarkable that the assumption of independence in the allosteric effects of three distinct regulatory elements accounts so well for the observations. In fact, to account for the broad range of $[\text{Ca}^{2+}]_i$ over which BK channels are regulated, independence of each of the Ca^{2+} -dependent regulatory elements is required. This definition of the principal regulatory elements by which Ca^{2+} controls activation of BK currents provides a critical step towards defining how the binding of Ca^{2+} may modulate channel activation. □

Methods

Generation and expression of mutant subunits

Disruption of the Ca^{2+} bowl region was accomplished with mutation of five consecutive aspartate residues to asparagines (5D5N; mSlo1 residues 898–902)⁹. Loci for mutational analysis of the RCK domain were guided by a comparison of Slo1 and Slo3 sequences in or near the region between β -sheet β A and helix α A (Fig. 1c) defined in the published crystal structure of the *Escherichia coli* K⁺ homologue¹². For structural comparisons, the Slo1 amino acid sequence was threaded onto the published structure of the RCK domain of the *E. coli* K⁺ channel¹². Because of the absence of residues in the linker between helix α A and sheet β B in the *E. coli* sequence, the position of residues in this linker shown in the diagram of the BK RCK structure (Fig. 1b) is uncertain.

Construction of mutations was accomplished using standard procedures²⁹. Methods of expression in *Xenopus* oocytes were as previously described³. In addition to mutations involving D362, D367, D369 and E399, a number of other single point mutations involving charged residues in the RCK domain (D481A, D482A, D434A and K448A) were also examined. Similar to wild-type Slo1 currents, each of these constructs exhibit a >250-mV negative shift in $V_{1/2}$ over Ca^{2+} from 0 to 100 mM, with little difference in the shape of the pCa versus $V_{1/2}$ relationship compared to control currents.

Physiological recordings

All recordings used inside-out patches and follow procedures in standard use in our laboratory^{3,30}. The pipette extracellular solution was 140 mM potassium methanesulphonate, 20 mM KOH, 10 mM HEPES, 2 mM MgCl_2 (pH 7.0). Test solutions bathing the cytoplasmic face of the patch membrane contained 140 mM potassium methanesulphonate, 20 mM KOH, 10 mM HEPES (pH 7.0), and one of the following:

5 mM EGTA (for nominally 0 Ca^{2+} , 0.5 and 1 μM Ca^{2+} solutions), 5 mM HEDTA (for 4 and 10 μM Ca^{2+} solutions), or no added Ca^{2+} buffer (for 30 μM and higher Ca^{2+}). The methods of calibration and preparation of Mg^{2+} solutions have been described previously³⁰. Excised patches were bathed with constantly flowing solutions of defined Ca^{2+} from a multibarrel local application system.

Conductance–voltage (G - V) curves were generated from tail currents following standard procedures⁵. Families of G - V curves represent averages from a set of patches at specific $[\text{Ca}^{2+}]_i$. For each patch, G - V curves were fit with

$$G(V) = G_{\max}/(1 + \exp^{(V-V_{1/2})/k}) \quad (1)$$

to provide estimates of $V_{1/2}$, the voltage of half activation, and k , the slope factor describing the voltage-dependence of the closed–open equilibrium. Estimates of the $V_{1/2}$ value for a given ionic condition represent the mean value for a set of five to ten individual patches. In all cases, error bars show s.e.m. In Fig. 3, for comparison of the magnitude of $V_{1/2}$ shifts, $V_{1/2}$ values for each construct were adjusted so that the values at 0 mM were identical. For all constructs, absolute $V_{1/2}$ values at 0 mM ranged from within about 20 mV, a range similar to that observed for different sets of patches from the same construct. Experiments were at room temperature (21–24 °C). Most salts and chemicals were from Sigma.

Allosteric models of BK current activation

For an allosteric model in which membrane voltage and one or more distinct ligand-binding sites all independently regulate the transition between closed and open conformations, the energy provided by each regulatory site is additive in activating the channel⁷. Such a model predicts that the shift in $V_{1/2}$ for a single divalent cation-binding site follows the following relationship for the case that there is one Ca^{2+} -binding site on each of the four BK α -subunits^{5,13}.

$$V_{1/2} = V_{1/2}(0) + \frac{4kT}{ze} \left\{ \ln \left[\frac{1 + [\text{Ca}^{2+}]_i/K_{C1}}{1 + [\text{Ca}^{2+}]_i/K_{O1}} \right] \right\}, \quad (2)$$

where $V_{1/2}(0)$ corresponds to activation $V_{1/2}$ at 0 Ca^{2+} , K_{C1} and K_{O1} correspond to the Ca^{2+} affinity of the closed and open channel, respectively, k is Boltzmann's constant and T is absolute temperature. For three distinct, independent regulatory elements, $V_{1/2}$ is predicted to shift in accordance with the following relationship:

$$V_{1/2} = V_{1/2}(0,0) + \frac{4kT}{ze} \left\{ \ln \left(\frac{1 + [\text{Ca}^{2+}]_i/K_{C1}}{1 + [\text{Ca}^{2+}]_i/K_{O1}} \right) + \ln \left(\frac{1 + [\text{Ca}^{2+}]_i/K_{C2}}{1 + [\text{Ca}^{2+}]_i/K_{O2}} \right) + \ln \left(\frac{1 + [\text{Ca}^{2+}]_i/K_{C3}}{1 + [\text{Ca}^{2+}]_i/K_{O3}} \right) \right\} \quad (3)$$

Thus, the shift in $V_{1/2}$ for each regulatory element is additive with the contributions of the other regulatory elements. Equation (3) was used to fit the relationship between $V_{1/2}$ and Ca^{2+} for all constructs (Fig. 4c). All points were fitted simultaneously. Estimates for Ca^{2+} binding affinities were: for the Ca^{2+} bowl, $K_{C1} = 4.5 \pm 1.7 \mu\text{M}$ and $K_{O1} = 2.0 \pm 0.7 \mu\text{M}$; for D362/D367, $K_{C2} = 17.2 \pm 4.0 \mu\text{M}$ and $K_{O2} = 4.6 \pm 1.0 \mu\text{M}$; and for E399, $K_{C3} = 4.1 \pm 3.5 \text{ mM}$ and $K_{O3} = 1.8 \pm 1.2 \text{ mM}$. With mutation to D362A/D367A, $K_{C2} = 6.6 \text{ mM}$ and $K_{O2} = 10.7 \text{ mM}$, accounting for the inhibitory effects of Ca^{2+} . Values above 10 mM were not included in the fit.

Received 16 April; accepted 19 June 2002; doi:10.1038/nature00956.

- Barrett, J. N., Magleby, K. L. & Pallotta, B. S. Properties of single calcium-activated potassium channels in cultured rat muscle. *J. Physiol. (Lond.)* **331**, 211–230 (1982).
- Moczydlowski, E. & Latorre, R. Gating kinetics of Ca^{2+} -activated K^{+} channels from rat muscle incorporated into planar lipid bilayers. Evidence for two voltage-dependent Ca^{2+} binding reactions. *J. Gen. Physiol.* **82**, 511–542 (1983).
- Cox, D. H., Cui, J. & Aldrich, R. W. Allosteric gating of a large conductance Ca-activated K^{+} channel. *J. Gen. Physiol.* **110**, 257–281 (1997).
- Cox, D. & Aldrich, R. Role of the β 1 subunit in large-conductance Ca^{2+} -activated K^{+} channel gating energetics. Mechanisms of enhanced Ca^{2+} sensitivity. *J. Gen. Physiol.* **116**, 411–432 (2000).
- Zhang, X., Solaro, C. & Lingle, C. Allosteric regulation of BK channel gating by Ca^{2+} and Mg^{2+} through a non-selective, low affinity divalent cation site. *J. Gen. Physiol.* **118**, 607–635 (2001).
- Cui, J., Cox, D. H. & Aldrich, R. W. Intrinsic voltage dependence and Ca^{2+} regulation of mslo large conductance Ca-activated K^{+} channels. *J. Gen. Physiol.* **109**, 647–673 (1997).
- Cui, J. & Aldrich, R. W. Allosteric linkage between voltage and Ca^{2+} -dependent activation of BK-type mslo1 K^{+} channels. *Biochemistry* **39**, 15612–15619 (2000).
- Horrigan, F. T., Cui, J. & Aldrich, R. W. Allosteric voltage gating of potassium channels I. Mslo ionic currents in the absence of Ca^{2+} . *J. Gen. Physiol.* **114**, 277–304 (2002).
- Schreiber, M. & Salkoff, L. A novel calcium-sensing domain in the BK channel. *Biophys. J.* **73**, 1355–1363 (1997).
- Schreiber, M., Yuan, A. & Salkoff, L. Transplantable sites confer calcium sensitivity to BK channels. *Nature Neurosci.* **2**, 416–421 (1999).
- Bian, S., Favre, I. & Moczydlowski, E. Ca^{2+} -binding activity of a COOH-terminal fragment of the *Drosophila* BK channel involved in Ca^{2+} -dependent activation. *Proc. Natl Acad. Sci. USA* **98**, 4776–4781 (2001).
- Jiang, Y., Pico, A., Cadene, M., Chait, B. T. & MacKinnon, R. Structure of the RCK domain from the *E. coli* K^{+} channel and demonstration of its presence in the human BK channel. *Neuron* **29**, 593–601 (2001).
- Shi, J. & Cui, J. Intracellular Mg^{2+} enhances the function of BK-type Ca^{2+} -activated K^{+} channels. *J. Gen. Physiol.* **118**, 589–606 (2001).
- Shao, L. R., Halvorsrud, R., Borg-Graham, L. & Storm, J. F. The role of BK-type Ca^{2+} -dependent K^{+} channels in spike broadening during repetitive firing in rat hippocampal pyramidal cells. *J. Physiol.* **521** (1), 135–146 (1999).
- Jones, E. M., Gray-Keller, M. & Fettiplace, R. The role of Ca^{2+} -activated K^{+} channel spliced variants in the tonotopic organization of the turtle cochlea. *J. Physiol.* **518**, 653–665 (1999).
- Adelman, J. P. et al. Calcium-activated potassium channels expressed from cloned complementary DNAs. *Neuron* **9**, 209–216 (1992).

17. Butler, A., Tsunoda, S., McCobb, D. P., Wei, A. & Salkoff, L. mSlo, a complex mouse gene encoding "maxi" calcium-activated potassium channels. *Science* **261**, 221–224 (1993).
18. Shen, K. Z. *et al.* Tetraethylammonium block of Slowpoke calcium-activated potassium channels expressed in *Xenopus* oocytes: evidence for tetrameric channel formation. *Pflügers Arch.* **426**, 440–445 (1994).
19. Meera, P., Wallner, M., Song, M. & Toro, L. Large conductance voltage- and calcium-dependent K⁺ channel, a distinct member of voltage-dependent ion channels with seven N-terminal transmembrane segments (S0–S6), an extracellular N terminus, and an intracellular (S9–S10) C terminus. *Proc. Natl Acad. Sci. USA* **94**, 14066–14071 (1997).
20. Wei, A., Solaro, C., Lingle, C. & Salkoff, L. Calcium sensitivity of BK-type KCa channels determined by a separable domain. *Neuron* **13**, 671–681 (1994).
21. Jiang, Y. *et al.* Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417**, 515–522 (2002).
22. Bellamacina, C. R. The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J.* **10**, 1257–1269 (1996).
23. Schreiber, M. *et al.* Slo3, a novel pH-sensitive K⁺ channel from mammalian spermatocytes. *J. Biol. Chem.* **273**, 3509–3516 (1998).
24. Moss, B. L. & Magleby, K. L. Gating and conductance properties of BK channels are modulated by the S9–S10 tail domain of the alpha subunit. A study of mSlo1 and mSlo3 wild-type and chimeric channels. *J. Gen. Physiol.* **118**, 711–734 (2001).
25. Golowasch, J., Kirkwood, A. & Miller, C. Allosteric effects of Mg²⁺ on the gating of Ca²⁺-activated K⁺ channels from mammalian skeletal muscle. *J. Exp. Biol.* **124**, 5–13 (1986).
26. Oberhauser, A., Alvarez, O. & Latorre, R. Activation by divalent cations of a Ca²⁺-activated K⁺ channel from skeletal muscle membrane. *J. Gen. Physiol.* **92**, 67–86 (1988).
27. Ferguson, W. B. Competitive Mg²⁺ block of a large-conductance, Ca²⁺-activated K⁺ channel in rat skeletal muscle. Ca²⁺, Sr²⁺, and Ni²⁺ also block. *J. Gen. Physiol.* **98**, 163–181 (1991).
28. Rothberg, B. S. & Magleby, K. L. Gating kinetics of single large-conductance Ca²⁺-activated K⁺ channels in high Ca²⁺ suggest a two-tiered allosteric gating mechanism. *J. Gen. Physiol.* **114**, 93–124 (1999).
29. Xia, X. M., Ding, J. P. & Lingle, C. J. Molecular basis for the inactivation of Ca²⁺- and voltage-dependent BK channels in adrenal chromaffin cells and rat insulinoma tumour cells. *J. Neurosci.* **19**, 5255–5264 (1999).
30. Lingle, C., Zeng, X.-H., Ding, J.-P. & Xia, X.-M. Inactivation of BK channels mediated by the N-terminus of the $\beta 3b$ auxiliary subunit involves a two-step mechanism: possible separation of binding and blockade. *J. Gen. Physiol.* **117**, 583–605 (2001).

Acknowledgements

We thank the members of our laboratory for encouragement and assistance during this work and L. Lavack for preparation and care of oocytes. We thank J. H. Steinbach, J. Nerbonne and L. Salkoff for discussions and comments on the manuscript. We also thank the Department of Anesthesiology, Washington University School of Medicine, for material support.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.J.L. (e-mail: cingle@morpheus.wustl.edu).

A bacteriolytic agent that detects and kills *Bacillus anthracis*

Raymond Schuch, Daniel Nelson & Vincent A. Fischetti

Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, New York 10021, USA

The dormant and durable spore form of *Bacillus anthracis* is an ideal biological weapon of mass destruction^{1,2}. Once inhaled, spores are transported by alveolar macrophages to lymph nodes surrounding the lungs, where they germinate; subsequent vegetative expansion causes an overwhelming flood of bacteria and toxins into the blood, killing up to 99% of untreated victims. Natural and genetically engineered antibiotic-resistant bacilli amplify the threat of spores being used as weapons, and heighten the need for improved treatments and spore-detection methods after an intentional release. We exploited the inherent binding specificity and lytic action of bacteriophage enzymes called lysins for the rapid detection and killing of *B. anthracis*. Here we show that the PlyG lysin, isolated from the γ phage of *B. anthracis*,

specifically kills *B. anthracis* isolates and other members of the *B. anthracis* 'cluster' of bacilli *in vitro* and *in vivo*. Both vegetative cells and germinating spores are susceptible. The lytic specificity of PlyG was also exploited as part of a rapid method for the identification of *B. anthracis*. We conclude that PlyG is a tool for the treatment and detection of *B. anthracis*.

Bacteriophage lysins are lytic agents used by double-stranded DNA (dsDNA) phages to coordinate bacterial host lysis with completion of viral assembly^{3,4}. Late in infection, lysin translocates from the cytoplasm into the cell-wall matrix, where it rapidly hydrolyses covalent bonds essential for peptidoglycan integrity, causing bacterial lysis and concomitant release of progeny phages. Lysin family members are often chimaeric proteins, consisting of a usually well-conserved catalytic domain fused to a largely divergent specificity or binding domain^{4,5}. High-affinity binding (the affinity constant $K_A = 3\text{--}6 \times 10^8$, similar to affinity-matured antibodies) is directed towards species- or strain-specific cell-wall carbohydrates that are often essential for viability, thus implying that intrinsic lysin resistance could be rare. Lysins have not previously been investigated as a means for bacterial control, although we have demonstrated efficient 'lysis from without' of pathogenic streptococci in the mouse nasopharynx using purified streptococcal phage lysins^{6,7}. On the basis of these findings, we thought that lysins might be able to deliver a rapid and specific lethal action to any particular bacterial pathogen, not just streptococci. The main requirement for developing lysins, as such, is a dsDNA phage specific for the pathogen of interest to serve as a lysin source. We investigated this possibility with a dsDNA phage specific for the biowarfare agent *B. anthracis*.

The dsDNA phages of *B. anthracis* form a homogeneous family⁸. We chose the γ phage as a lysin source because of its specificity for *B. anthracis* and its usage by the US Centers for Disease Control and Prevention (CDC, Atlanta, Georgia) for identification of this pathogen^{2,9}. It is known that γ phages infect most *B. anthracis* isolates, including some rare *Bacillus cereus* strains that could represent *B. anthracis* cured of its virulence plasmid or an environmental reservoir of potential progenitors¹⁰. Isolates of RSVF1 (streptomycin-resistant *B. cereus* strain 4342 from the American Type Culture Collection, ATCC) and *B. anthracis* are monomorphic at multiple allozyme loci, and therefore are part of the same highly related cluster of isolates within the *B. cereus* lineage¹¹. Consistent with this, we found that RSVF1 was sensitive to the γ phage (Table 1) and displayed several other features typical of *B. anthracis*: matt colony morphology, filamentous structure, lack of motility, and characteristic sequences in the hypervariable *vrroA* locus (data not shown). For these reasons, RSVF1 is an appropriate representative of the γ -phage-sensitive *B. anthracis* cluster of *B. cereus* for use in this study.

An expression library of γ phage proteins was screened for agents capable of lysing RSVF1 'from without'. Each lytic clone identified contained a 702-base-pair (bp) γ open reading frame (ORF) encoding a product that is homologous to *N*-acetylmuramoyl-L-alanine amidases (amidases), a class of phage lysins (Fig. 1a). The homology was restricted to the catalytic amino-terminal halves, and absent in the cell-wall-specific carboxy-terminal binding domains^{4,5}. Recombinant γ lysin (called PlyG, for phage lysin γ) was purified to homogeneity by column chromatography (Fig. 1b). Gel filtration confirmed a predicted relative molecular mass of about 27,000 (M_r 27K), and suggested that PlyG acts as a monomer and is not proteolytically processed.

To evaluate activity and specificity, 0.5 U of PlyG was initially added in drops to bacterial lawns of RSVF1 (Fig. 1c) and isolates of *B. anthracis* gathered worldwide as well as other bacilli from the *B. cereus* lineage (*B. cereus* and *Bacillus thuringiensis*) and other Gram-positive genera (Table 1). RSVF1 was the only *B. cereus* strain found to be as sensitive to PlyG killing (and sensitive to the γ phage) as the diverse set of *B. anthracis* isolates (Table 1). *B. cereus* ATCC 10987, a strain closely related to *B. anthracis*¹¹, was slightly susceptible to

Table 1 Susceptibility to γ phage infection and purified PlyG activity

Isolate or strain*	γ phage titre (PFU ml ⁻¹)†	γ lysin activity‡	Strain source
<i>B. anthracis</i> §			
Vollum	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
Ames	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A1.a/10 (USA)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A1.b/23 (Turkey)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A2/29 (Pakistan)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A3.a/34 (Korea)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A3.b/57 (China)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
A4/69 (Pakistan)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
B/80 (France)	10 ⁸ –10 ¹⁰	+++	L. W. Mayer
Δ Sterne	10 ⁸ –10 ¹⁰	+++	A. L. Turetsky
VNR1 Δ 1	10 ⁸ –10 ¹⁰	+++	A. L. Turetsky
Δ Ames	10 ⁸ –10 ¹⁰	+++	A. L. Turetsky
NNR1 Δ 1	10 ⁸ –10 ¹⁰	+++	A. L. Turetsky
Δ NH1	10 ⁸ –10 ¹⁰	+++	A. L. Turetsky
<i>B. cereus</i>			
RSVF1	2.2 $\times$ 10 ¹⁰	+++	ATCC
ATCC 10987	<10	+/-	ATCC
ATCC 27348	<10	-	A. Aronson
T	<10	-	A. Aronson
ATCC 14579	<10	-	A. Keynan
ATCC 15816	<10	-	ATCC
RTS134	<10	-	A. Keynan
RSVF2	<10	-	A. Keynan
ATCC 13472	<10	-	H.-W. Ackerman
S154-2 HER1414	<10	-	H.-W. Ackerman
<i>B. thuringiensis</i>			
T24	<10	-	H.-W. Ackerman
ATCC 33679	<10	-	A. Aronson
ATCC 35866	<10	-	R. McNall

*Additional bacteria were also tested and were resistant to the γ phage and lysin. These include *B. subtilis*, *B. megaterium*, *B. stearothermophilus*, *B. pumilus*, *B. brevis*, *B. licheniformis*, *Brevibacillus laterosporus*, *Sporosarcina ureae*, *Streptococcus pneumoniae*, *Pseudomonas aeruginosa* and *Escherichia coli*.

†Titres indicate the maximum plaque-forming units per ml of donor phage stock obtained with the indicated bacterium.

‡Designations refer to the degree of bacterial lawn clearance mediated by purified lysin and are as follows: +++, complete clearance at the site of lysin application; +/-, partial clearance observable only at the centre of lysin application; -, no clearance.

§Numerical designations for *B. anthracis* strains represent distinct genotypes by MLVA (multiple-locus variable number of tandem repeat analysis)³⁰. The geographical region of isolation for some *B. anthracis* strains is indicated.

PlyG, but all other strains examined were resistant. A more sensitive test of PlyG-mediated killing was evaluated in buffer containing 20 U of PlyG. After 15 min, RSVF1 was nearly sterilized (Fig. 2a), whereas the viability of ATCC 10987 was reduced about 100-fold. Other strains examined were largely resistant, even after a 3-h incubation. PlyG can clearly direct a potent and specific lethal action to the *B. anthracis* cluster, exhibiting a substrate specificity that closely matches the γ phage host range. Moreover, the capsulated state of several *B. anthracis* strains examined indicated that the capsule does not block access of PlyG to the cell wall.

We found that PlyG, like most lysins, is a highly active enzyme. The addition of 2 U of PlyG to 1.0×10^4 RSVF1 caused a pronounced release of intracellular ATP within 10 s (measured as light emitted by luciferin/luciferase) (Fig. 2b). This rapid lytic effect was not observed with other isolates tested, suggesting that the ATP release assay is a rapid diagnostic tool for γ -sensitive bacilli. In a separate kinetic analysis of RSVF1 viability after PlyG treatment, as little as 2 U added to 1 ml of log-phase cells effected a 17,000-fold decrease within 20 s, and near sterilization at 2 min (Fig. 2c). The loss of bacterial absorbance lagged a few minutes behind the loss in viability, implying that PlyG does not cause immediate cell-wall degradation. To visualize the lytic effect, we used phase-contrast microscopy of PlyG-treated RSVF1. The normally filamentous RSVF1 (Fig. 3a) rapidly converts to short rod- and minicell-like forms 30 s after exposure (Fig. 3b); nearly complete loss of cytoplasmic material occurs by 15 min, leaving 'ghost' cells (Fig. 3c). Transmission electron micrographs of the rod form reveals cytoplasmic membrane extrusions from distinct regions of cell-wall hydrolysis at polar and septal positions (Fig. 3d), and the formation of ghost-like forms (Fig. 3e).

We next tested the effect of PlyG on RSVF1 using a mouse model of infection. The intraperitoneal (i.p.) injection of some *B. cereus* isolates can induce a rapidly fatal illness^{12,13}. We injected about 1.0×10^6 RSVF1 cells and observed a characteristic progression of symptoms in BALB/c mice, leading to death in 5 h or less (Fig. 4). At death, many mice exhibited severe oedema at the inoculation site, and haemorrhaging through the eyes and mouth. When PlyG (50 U in 0.5 ml) was injected i.p. 15 min after infection, a pronounced therapeutic effect was observed: 13 of 19 mice (68.4%) recovered fully, and the remaining animals survived 6–21 h. When 150 U of PlyG were used in 0.5 ml, a similar rate of recovery was observed (76.9%), possibly owing to the convoluted nature of the peritoneum and the inability of the enzyme to reach some bacteria concealed within peritoneal folds. No toxicity was detected upon administration of PlyG alone. PlyG does, therefore, rapidly kill γ -sensitive bacteria in an infected animal.

To evaluate the therapeutic potential of PlyG, we looked for resistance to its binding or catalytic activity. Repeated exposure of RSVF1 to low or high PlyG concentrations on either agar plates or in liquid cultures did not identify spontaneously resistant mutants (a frequency of $<5.0 \times 10^{-9}$), although we readily identified mutants resistant to novobiocin and streptomycin (Table 2) using a similar strategy. We then subjected RSVF1 to mutagenesis with methanesulphonic acid ethyl ester (EMS) and observed roughly 1,000-fold and 10,000-fold increases in novobiocin and streptomycin resist-

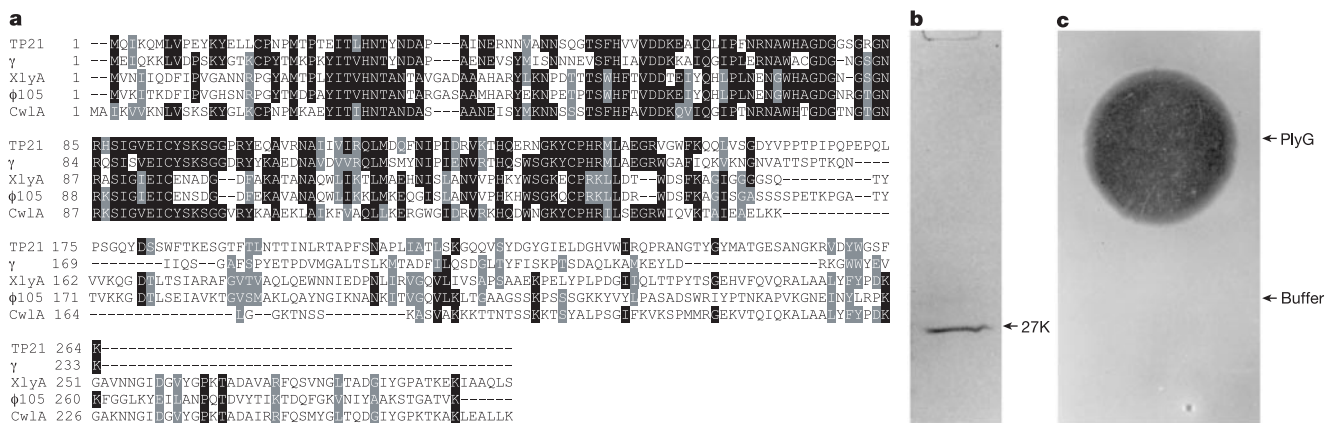
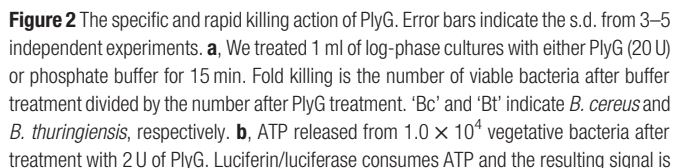


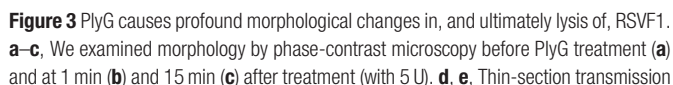
Figure 1 Characterization of PlyG. **a**, Sequence alignment of the γ lysin with *B. cereus* and *B. subtilis* phage amidases (TP21 and ϕ 105, respectively) and *B. cereus* and *B. subtilis* prophage amidases (CwIA and XlyA, respectively). Dark shading represents sequence identity and lighter shading represents similarity. **b**, SDS-polyacrylamide gel

electrophoresis of purified PlyG stained with Coomassie blue. The molecular mass was estimated with Kaleidoscope (Bio-Rad) standards that are not shown. **c**, An RSVF1 lawn on a BHI plate treated with 0.5 U of purified PlyG. The clearing zone is 18.0 mm in diameter.



We next examined the ability of PlyG to degrade germinating spores. In the dormant state, spore peptidoglycan, or cortex, is protected from lysozymes and amidases by a proteinaceous coat¹⁵⁻¹⁷; however, within 10 min of inducing germination, coat porosity increases¹⁸. We therefore reasoned that the subjacent peptidoglycan may be rendered susceptible to PlyG after the induction of germination. To evaluate this, we induced 1-ml aliquots of about

10^8 heat-activated spores from RSVF1, closely related *B. cereus* (ATCC 14579) and *B. thuringiensis* (ATCC 33679) strains, and *Bacillus subtilis* to germinate for 5 min with L-alanine (or treated for 5 min with D-alanine, a germination inhibitor). The suspensions were treated with PlyG (10 U) for 5 min, washed and the resulting viabilities were determined (Fig. 5a). The D-alanine-treated spores were resistant to PlyG, suggesting that the enzyme is not being carried by spores onto agar plates for determinations of colony-forming units. Among the L-alanine-treated spores, only RSVF1 was sensitive after the induction of germination, showing a 7,500-fold decrease in viability. A bacteriocidal action, therefore, occurs rapidly after the induction of germination, probably when PlyG can access the cortex.



electron micrographs (45,000 $\times$) 1 min (**d**) and 10 min (**e**) after treatment (with 5 U of PlyG). The arrows indicate membrane protrusions, and remnants thereof, pressing through zones of peptidoglycan degradation.

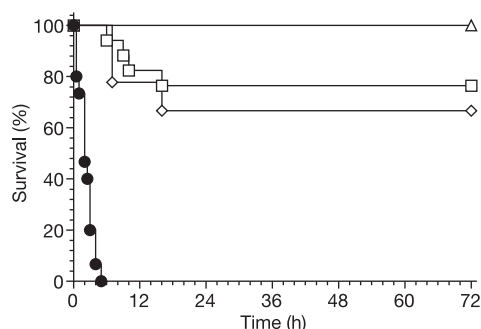


Figure 4 Survival of PlyG-treated BALB/c mice infected with RSVF1. Mice were injected i.p. with $\sim 1.0 \times 10^6$ RSVF1 colony-forming units and treated after 15 min with either phosphate buffer (circles, $n = 15$), 50 U PlyG (diamonds, $n = 19$), or 150 U PlyG (squares, $n = 13$). As a control for toxicity, mice were injected with 50 U PlyG (triangles, $n = 5$) alone. The experiment was terminated after 72 h. Administration of 50 U and 150 U to the infected mice was significantly protective compared with the buffer control ($P < 0.0001$). The median survival time for buffer-treated mice was 2 h.

The ability of PlyG to kill germinating spores was exploited to develop a rapid and specific system for detecting γ -sensitive spores using a hand-held luminometer. Spores were immobilized on filters and incubated in consecutive 5-min rounds with germinant and PlyG (2 U); ATP release from degrading spores was then measured as light emitted in the presence of luciferin/luciferase. Using 2.5×10^3 spores, we detected a PlyG-mediated signal only with germinating RSVF1 spores (Fig. 5b), demonstrating both the recognition specificity of PlyG and its rapid lytic action. This signal was detected 10 min after germinant addition and only 5 min after PlyG was added. With mixed spore preparations, only a combination containing RSVF1 yielded a light signal (Fig. 5c); spore cortex fragments released from the germinating spores of related *Bacillus* isolates in the mix do not competitively inhibit PlyG. The sensitivity of this system was demonstrated using as few as about 100 RSVF1 spores, which yielded a signal 60 min after the addition of PlyG (Fig. 5d). No signal was detected in the presence of other germinating spore types, and was, therefore, specific to the γ -sensitive spores. This sensitivity, combined with the specificity, rapidity and portable nature of our detection method, suggests applications in monitoring domestic and battlefield use of *B. anthracis*.

In this study, we identified the first lysin (to our knowledge) from a *B. anthracis* bacteriophage and demonstrated its potent lytic effect

Table 2 Intrinsic PlyG resistance in wild-type and mutagenized RSVF1

Treatment	RSVF1		RSVF1 + EMS	
	Resistance	Frequency	Resistance	Frequency
Novobiocin $3.5 \mu\text{g ml}^{-1}$	Yes	2.4×10^{-9}	Yes	2.1×10^{-6}
Streptomycin $150 \mu\text{g ml}^{-1}$	Yes	4.0×10^{-10}	Yes	4.3×10^{-6}
PlyG 20 U	No	$<5.0 \times 10^{-9}$	No	$<5.0 \times 10^{-9}$
PlyG 0.5–50 U	No	$<5.0 \times 10^{-9}$	No	$<5.0 \times 10^{-9}$
PlyG 20 U*	No	$<5.0 \times 10^{-9}$	No	$<5.0 \times 10^{-9}$

Descriptions of each treatment are provided in the Methods. PlyG 20 U and PlyG 0.5–50 U refer to the amounts used over a 4-day period. PlyG 20 U* refers to the amount present in a 24-h culture, before plating.

on *B. anthracis* and other members of the *B. anthracis* cluster of *B. cereus*. Our findings encourage the continued development of PlyG as a means to prevent or treat anthrax and as a tool to detect vegetative or spore forms of *B. anthracis*. In the face of new and emerging bacterial pathogens, the widespread distribution of antibiotic resistance, and the real threat of bioterrorism and biowarfare, new adjuncts or alternatives to antibiotics, such as lysins, are imperative. In addition to their rapid and specific lytic activities, lysins are attractive compounds for development for several reasons. First, the biochemical pathways leading to the species- and serovar-specific cell-wall antigens bound by lysins are largely new or under-used targets in current antimicrobial therapies for Gram-positive bacteria. Second, the modular design of lysins, with their independent functional domains, makes them ideal for domain-swapping studies in which bacterial specificities and catalytic activities are improved or adapted for use with other pathogens⁴. Third, because the catalytic and binding targets of lysins are largely essential for viability, lysin resistance should be rare. Our inability to detect resistance to PlyG certainly suggests that its peptidoglycan catalytic target and carbohydrate-binding sites cannot be easily modified by the bacillus to prevent lysin action. Finally, we must realize that lysins are highly evolved enzymes, modified and improved for high activity and specificity over the millennia. Considering the staggering global population of dsDNA phages (estimated to be more than 10^{30})¹⁹, lysins are already one of the most ubiquitous and successful antimicrobial agents on Earth. Bacteriophages, and their corresponding lytic activities, represent an enormous untapped pool of agents with which to control human pathogens. □

Methods

Strains

Most strains were grown at 30 °C in Luria broth (LB) or brain–heart infusion broth (BHI). *B. anthracis* was grown on sheep blood agar (trypticase soy base) with or without a CO₂

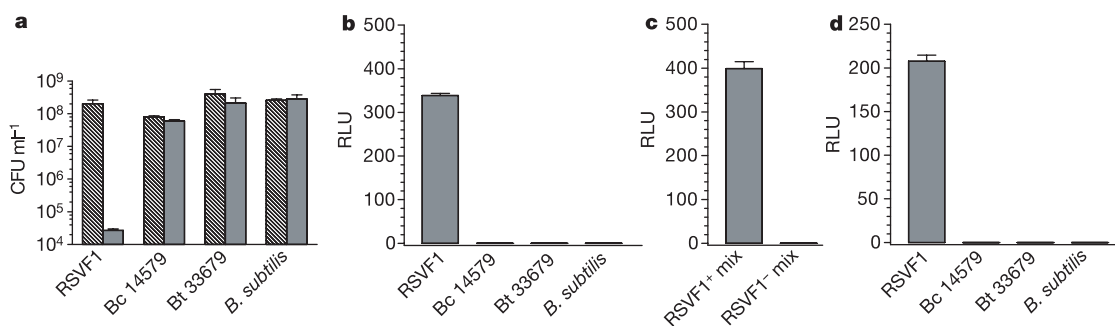


Figure 5 PlyG-mediated spore killing and detection. Error bars indicate s.d. from 3–5 independent determinations. **a**, The effect of PlyG on spore viability. Spores were incubated with D-alanine (hatched bars) or L-alanine (grey bars) and treated with PlyG (10 U). The resulting viability is shown. **b**, Specific detection of 2.5×10^3 germinating spores. ATP released 5 min after PlyG treatment (2 U) generates the indicated light signal.

c, Detection of RSVF1 in a spore mixture. Mixtures containing 2.5×10^3 spores each of *B. cereus* (Bc) 14579, *B. thuringiensis* (Bt) 33679 and *B. subtilis* with (RSVF1⁺) or without (RSVF1⁻) RSVF1 were induced to germinate; the light release 5 min after PlyG treatment (2 U) is indicated. **d**, One hundred spores of the indicated types was induced to germinate; the light release 60 min after addition of 2 U PlyG is shown.

atmosphere. Work with *E. coli* XLI-Blue (Stratagene) and *B. anthracis* was performed at 37 °C, whereas *B. stearothermophilis* was handled at 55 °C. RSVF1, except for the absence of virulence plasmids, is nearly indistinguishable from *B. anthracis*^{20–22}. The *vrnA* locus and motility of RSVF1 were analysed as described^{23,24}. The susceptibility of *B. anthracis* to PlyG was examined by L. W. Mayer (CDC) and Abraham L. Turetsky (Aberdeen Proving Grounds, Maryland).

Virus manipulations

The γ phage was obtained from H.-W. Ackermann. Susceptibilities were initially tested using drops of $\sim 10^7$ phages applied to fresh lawns of the indicated strains. A phage stock containing 2.2×10^{10} plaque-forming units (PFU) per ml was prepared on RSVF1 as described⁶, and used for the titre determinations in Table 1.

Identification of γ lysis

We partially digested 5- μ g aliquots of γ DNA with *Tsp*509I and cloned 0.5–3.0-kilobase fragments into the arabinose-inducible expression vector pBAD24 (ref. 25). The library was transformed into *E. coli* and screened for lysis activity in a manner derived from another protocol²⁶. The expression library was replica plated on to glass plates with LB containing 0.25% L-arabinose and 100 μ g ml⁻¹ ampicillin. After overnight incubation at 37 °C, the plates were inverted over lids containing 5 ml of chloroform to permeabilize the library. After 5 min, the library was overlain with 3 ml of 0.75% LB soft agar containing 300 μ l of a tenfold-concentrated log-phase culture of RSVF1 and incubated for 24 h. Clear lytic zones appeared at the periphery of lysis-encoding library members.

Protein purification

PlyG was induced from XLI-Blue/pBAD24::plyG with 0.25% L-arabinose in overnight LB cultures. Cells were washed, resuspended in 50 mM Tris buffer at pH 8.0, and lysed with chloroform to yield crude PlyG. PlyG, which passed through a HiTrap Q Sepharose XL column (Amersham Bioscience), bound to a Mono S HR 5/5 column (Amersham Biosciences) and was eluted in a linear gradient containing 1 M NaCl. Analysis of the N-terminal protein sequence confirmed the identity of PlyG.

In vitro lysis activity

Activity was examined in several ways. A Spectramax Plus 384 spectrophotometer (Molecular Devices) followed the drop in absorbance at 600 nm (A_{600}) of log-phase RSVF1 incubated for 15 min at 37 °C with PlyG. Enzyme activity in units was determined as described⁷, using lysis of exponentially growing RSVF1 with serial dilutions of PlyG. Fortunately, 1 U corresponds to 1 μ g of PlyG. For the liquid killing assay, 1.0 ml of log-phase cells was treated with the indicated amounts of PlyG at 37 °C and, at the specific time points, samples were washed to remove lysis and plated for enumeration. The ATP release assay was performed using the PROFILE-1 reagent kit and model 3550i microluminometer (New Horizons Diagnostics). Use of this system, and detailed validation studies, have been described²⁷. In brief, filter-immobilized vegetative bacteria were treated with 2 U of PlyG in 150 μ l for 2 min; a luciferin/luciferase reagent (50 μ l) was then added and light release was assayed in a 10-s integration. Relative light units (RLU) produced by RSVF1 were consistently 10–20% of total releasable light, determined using a detergent mix that releases all intracellular ATP. Total RLU was similar for each sample. The correlation between RSVF1 concentration and RLU was performed as described looking at total ATP release²⁷.

Mutagenesis and screening for PlyG resistance

RSVF1 was treated with EMS as described²⁸. Mutagenized cells were incubated with 20 U PlyG for 30 min at 37 °C, washed, and either plated or incubated overnight in BHI liquid. Colonies arising from plated cells were evaluated spectrophotometrically for resistance to 20 U of PlyG. For overnight BHI cultures, log-phase cells were established and treated with PlyG as above; this sequence was repeated over 4 d (using either 20 U PlyG at each treatment or increasing amounts of PlyG (0.5–50 U)).

Spontaneous lysis resistance was examined as described⁶. Spontaneous and EMS-induced mutation frequencies were estimated by determining resistance to 150 μ g ml⁻¹ streptomycin or to 3.5 μ g ml⁻¹ novobiocin.

Microscopy

We treated 1 ml of log-phase RSVF1 with PlyG (5 U) for the indicated periods. For phase-contrast microscopy, cells were viewed with an Eclipse E400 microscope (Nikon). For electron microscopy, cells were examined as described⁶.

In vivo lysis activity

Female BALB/c mice 4–8 weeks old were infected as previously described^{12,13}. We i.p. injected 0.1-ml aliquots of 1.0×10^6 log-phase RSVF1 in 50 mM potassium phosphate buffer at pH 7.4. After 15 min, 0.5 ml of buffer or PlyG in buffer were injected i.p. Injections of PlyG alone (no bacteria) were also performed to assess toxicity. Mice were then monitored for 3–4 d.

Spore preparation and germination

Samples containing 95–99% spores were prepared as described²⁹. Aliquots of 2.0×10^8 spores were heat activated at 65 °C for 5 min and suspended in 1.0 ml tryptic soy broth (TSB) with 100 mM L-alanine or D-alanine for 5 min at 37 °C. Samples were then treated with 10 U of PlyG for 5 min, washed three times (to remove residual PlyG), and plated for counting. TSB with L-alanine induced >99% germination, whereas D-alanine blocked germination.

For spore detection, we modified the germinating spore-killing protocol for use with the model 3550i microluminometer. Filter-immobilized, heat-activated spores were treated with 100 μ l TSB and 100 mM L-alanine for 5 min, and 150 μ l PlyG (2 U) for a further 5 min. Luciferin/luciferase (50 μ l) was added and RLU were determined at the indicated time point. When PlyG or the germinant was omitted, no light signal was detected. When PlyG was replaced with a strong detergent mix, and total ATP was released, similar RLU were observed for equivalent numbers of each germinating spore type.

Received 5 June; accepted 22 July 2002; doi:10.1038/nature01026.

- Mock, M. & Fouet, A. Anthrax. *Annu. Rev. Microbiol.* **55**, 647–671 (2001).
- Inglisby, T. V. *et al.* Anthrax as a biological weapon, 2002: Updated recommendations for management. *J. Am. Med. Assoc.* **287**, 2236–2252 (2002).
- Wang, I. N., Smith, D. L. & Young, R. Holins: The protein clocks of bacteriophage infections. *Annu. Rev. Microbiol.* **54**, 799–825 (2000).
- Lopez, R., Garcia, E., Garcia, P. & Garcia, J. L. The pneumococcal cell wall degrading enzymes: A modular design to create new lysins? *Microb. Drug Resist.* **3**, 199–211 (1997).
- Loessner, M. J., Kramer, K., Ebel, F. & Scherer, S. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol. Microbiol.* **44**, 335–349 (2002).
- Loeffler, J. M., Nelson, D. & Fischetti, V. A. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science* **294**, 2170–2172 (2001).
- Nelson, D., Loomis, L. & Fischetti, V. A. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc. Natl Acad. Sci. USA* **98**, 4107–4112 (2001).
- Redmond, C., Henderson, I., Turnbull, P. C. B. & Bowen, J. Phage from different strains of *Bacillus anthracis*. *Salisbury Med. Bull. Spec. Suppl.* **87**, 60–63 (1996).
- Brown, E. R. & Cherry, W. B. Specific identification of *Bacillus anthracis* by means of a variant bacteriophage. *J. Infect. Dis.* **96**, 34–39 (1955).
- Turnbull, P. C. B. Definitive identification of *Bacillus anthracis*—a review. *J. Appl. Microbiol.* **87**, 237–240 (1999).
- Helgason, E. *et al.* *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*—one species on the basis of genetic evidence. *Appl. Environ. Microbiol.* **66**, 2627–2630 (2000).
- Burdon, K. L., Davis, J. S. & Wende, R. D. Experimental infection of mice with *Bacillus cereus*: Studies of pathogenesis and pathologic changes. *J. Infect. Dis.* **117**, 307–316 (1967).
- Lamanna, C. & Jones, L. Lethality for mice of vegetative and spore forms of *Bacillus cereus* and *Bacillus cereus*-like insect pathogens injected intraperitoneally and subcutaneously. *J. Bacteriol.* **85**, 532–535 (1963).
- Fischetti, V. A. Phage antibacterials make a comeback. *Nature Biotechnol.* **19**, 734–735 (2001).
- Makino, S., Ito, N., Inoue, T., Miyata, S. & Moriyama, R. A spore-lytic enzyme released from *Bacillus cereus* spores during germination. *Microbiology* **140**, 1403–1410 (1994).
- Jenkinson, H. F., Kay, D. & Mandelstam, J. Temporal dissociation of late events in *Bacillus subtilis* sporulation from expression of genes that determine them. *J. Bacteriol.* **141**, 793–805 (1980).
- Henriques, A. O. & Moran, C. P. Jr. Structure and assembly of the bacterial endospore coat. *Methods* **20**, 95–110 (2000).
- Santo, L. Y. & Doi, R. H. Ultrastructural analysis during germination and outgrowth of *Bacillus subtilis* spores. *J. Bacteriol.* **120**, 475–481 (1974).
- Brussow, H. & Hendrix, R. W. Phage genomics: Small is beautiful. *Cell* **108**, 13–16 (2002).
- Helgason, E., Caugant, D. A., Olsen, I. & Kolsto, A. B. Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections. *J. Clin. Microbiol.* **38**, 1615–1622 (2000).
- Pannucci, J., Okinaka, R. T., Sabin, R. & Kuske, C. R. *Bacillus anthracis* pXO1 plasmid sequence conservation among closely related bacterial species. *J. Bacteriol.* **184**, 134–141 (2002).
- Ticknor, L. O. *et al.* Fluorescent amplified fragment length polymorphism analysis of norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. *Appl. Environ. Microbiol.* **67**, 4863–4873 (2001).
- Schmitt, C. K. *et al.* Absence of all components of the flagellar export and synthesis machinery differentially alters virulence of *Salmonella enterica* serovar Typhimurium in models of typhoid fever, survival in macrophages, tissue culture invasiveness, and calf enterocolitis. *Infect. Immun.* **69**, 5619–5625 (2001).
- Jackson, P. J. *et al.* Characterization of the variable-number tandem repeats in *vrnA* from different *Bacillus anthracis* isolates. *Appl. Environ. Microbiol.* **63**, 1400–1405 (1997).
- Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
- Loessner, M. J., Wendlinger, G. & Scherer, S. Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: A new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. *Mol. Microbiol.* **16**, 1231–1241 (1995).
- Stopa, P. J., Tieman, D., Coon, P. A., Milton, M. M. & Paterno, D. Detection of biological aerosols by luminescence techniques. *Field Anal. Chem. Technol.* **3**, 283–290 (1999).
- Corran, J. The induction of supersuppressor mutants of *Bacillus subtilis* by ethyl methanesulphonate and the posttreatment modification of mutation yield. *Mol. Gen. Genet.* **103**, 42–57 (1968).
- Mazas, M., Martinez, S., Lopez, M., Alvarez, A. B. & Martin, R. Thermal inactivation of *Bacillus cereus* spores affected by the solutes used to control water activity of the heating medium. *Int. J. Food Microbiol.* **53**, 61–67 (1999).
- Keim, P. *et al.* Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol.* **182**, 2928–2936 (2000).

Acknowledgements

We thank P. J. Piggot, L. W. Mayer, A. L. Turetsky, A. Aronson, A. Keynan, H.-W. Ackerman, R. J. McNall and T. A. Kokjohn for their gifts of strains; E. Spicas at the Bio-imaging Resource Center at The Rockefeller University for help with electron microscopy; R. L. Russell and S. Zhu for technical help; New Horizons Diagnostics for their luminometer and reagents and P. Model and members of the V.A.F. laboratory for

reviewing this manuscript. We also thank A. Keynan for reviewing this manuscript and for advice regarding spore preparation and germination. This work was supported by a grant from the Defense Advanced Research Projects Agency (DARPA).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.A.F. (e-mail: vaf@rockefeller.edu). The γ phage *plyG* sequence has been deposited in GenBank (accession number AF536823).

A saponin-detoxifying enzyme mediates suppression of plant defences

K. Bouarab, R. Melton, J. Peart, D. Baulcombe & A. Osbourn

Sainsbury Laboratory, John Innes Centre, Norwich NR4 7UH, UK

Plant disease resistance can be conferred by constitutive features such as structural barriers or preformed antimicrobial secondary metabolites. Additional defence mechanisms are activated in response to pathogen attack and include localized cell death (the hypersensitive response)^{1,2}. Pathogens use different strategies to counter constitutive and induced plant defences, including degradation of preformed antimicrobial compounds³ and the production of molecules that suppress induced plant defences^{4–6}. Here we present evidence for a two-component process in which a fungal pathogen subverts the preformed antimicrobial compounds of its host and uses them to interfere with induced defence responses. Antimicrobial saponins are first hydrolysed by a fungal saponin-detoxifying enzyme. The degradation product of this hydrolysis then suppresses induced defence responses by interfering with fundamental signal transduction processes leading to disease resistance.

Many solanaceous plants produce glycosylated steroidal and/or steroidal alkaloid saponins⁷. In *Lycopersicon* species such as tomato (*L. esculentum*), the main saponin is the steroidal glycoalkaloid α -tomatine, which has potent antifungal activity^{7,8}. The tomato leaf spot fungus *Septoria lycopersici* produces tomatinase, an extracellular enzyme that hydrolyses glucose from α -tomatine to give β -2-tomatine, which is substantially less inhibitory to fungal growth^{9–15}. Targeted mutation of the tomatinase gene results in loss of ability to degrade α -tomatine and enhanced sensitivity to α -tomatine^{12,16}. Tomatinase-deficient mutants of *S. lycopersici* are not obviously compromised in their ability to cause disease on tomato leaves but trigger enhanced cell death and elevated expression of plant defence genes in the early stages of infection¹². These effects may be due to subtle differences in growth of the mutant and wild-type strains during infection. Alternatively, they may be an indication that tomatinase interferes with mechanisms of disease resistance in plants¹². This led us to investigate whether *S. lycopersici* can infect and cause disease on another solanaceous species, *Nicotiana benthamiana*. Virus-induced gene silencing (VIGS) methodology has been developed for *N. benthamiana*, which makes this species an attractive model for further analysis¹⁷.

We inoculated three strains of *S. lycopersici* (NEV, 16R and NY) onto *N. benthamiana* leaves. All three strains caused spreading disease lesions and extensive tissue damage (shown for NEV in Fig. 1). The fungus infected the leaves through the stomata (Fig. 1g) and grew intercellularly in the mesophyll tissue, forming V-shaped branches around the plant cells as it does in tomato (ref. 12 and Fig.

1h, i). By contrast, tomatinase-deficient mutants of NEV failed to cause disease on *N. benthamiana* leaves (Fig. 1c, f). Microscopic analysis showed intense trypan blue staining of the guard cells of challenged leaves (Fig. 1j) and evidence of localized cell death in the mesophyll tissue at the infection site (Fig. 1k). This cell death was first apparent 2 d after inoculation, which is consistent with the hypersensitive response.

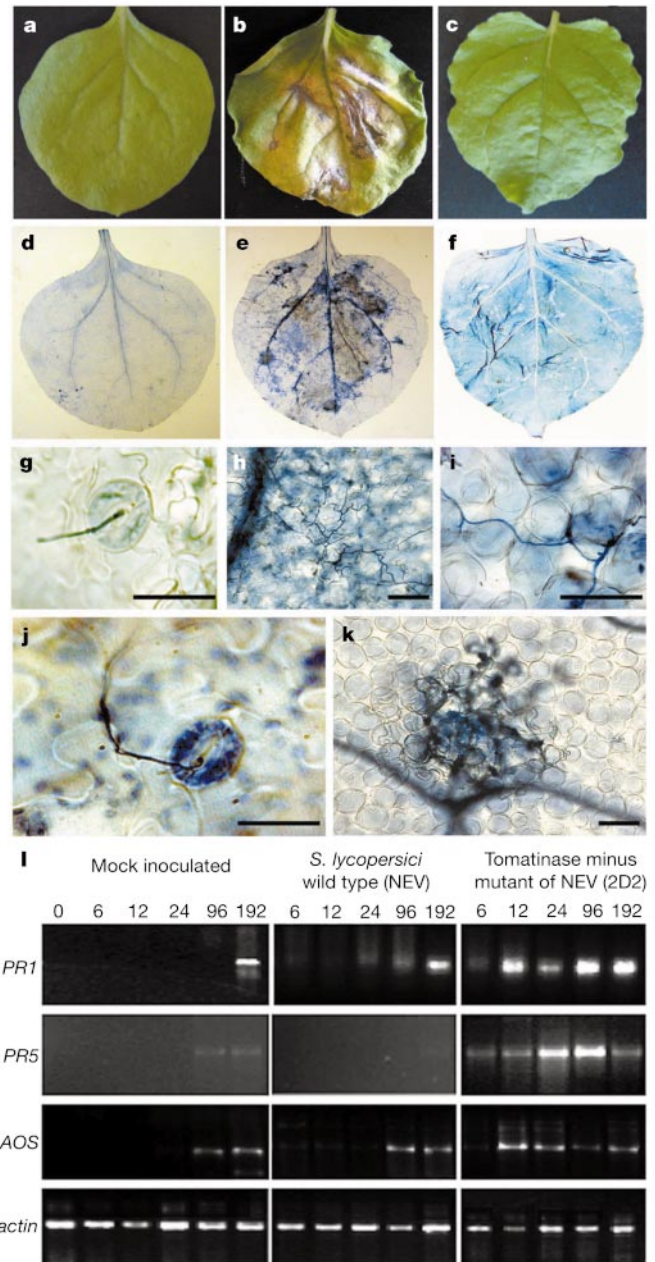


Figure 1 Infection of *N. benthamiana* by *S. lycopersici* strains. **a–c**, Leaves from *N. benthamiana* either mock inoculated (**a**) or inoculated with the wild-type *S. lycopersici* strain NEV (**b**) or the 2D2 tomatinase-deficient mutant of NEV (**c**). **d–f**, Leaves equivalent to those shown in **a–c** stained with lactophenol and trypan blue (**d**, mock inoculated; **e**, NEV wild-type strain; **f**, 2D2). **g–k**, Microscopic analysis of infection by the NEV wild-type strain (**g–i**) and 2D2 (**j, k**). Scale bars, 40 μ m. **l**, Induction of defence-related gene expression by wild-type and tomatinase-deficient strains of *S. lycopersici*. RNA was extracted from *N. benthamiana* leaves that had been either mock inoculated or inoculated with spores of the wild-type *S. lycopersici* strain NEV or the tomatinase-deficient mutant of NEV, 2D2. Expression of the defence-related genes *PR1*, *PR5* and *AOS*, and *actin* as a control was analysed by RT-PCR. Results were reproduced in three independent experiments. Numbers along the top of the panel indicate time (h) after inoculation.

Experiments with three independently derived tomatinase gene replacement mutants and two ectopic transformants verified that the effects were due specifically to mutation of the tomatinase gene and not to mutations elsewhere in the genome (data not shown). Polymerase chain reaction with reverse transcription (RT-PCR) experiments confirmed that tomatinase-deficient mutants triggered elevated expression of defence-related genes^{18–21} early in infection (Fig. 1l). Thus, *N. benthamiana* can serve as a host for *S. lycopersici* and this infection process requires tomatinase. In the absence of this enzyme, the fungus cannot proliferate in the leaf tissue and triggers localized cell death.

To investigate the nature of the induced defence response of *N. benthamiana* to tomatinase-deficient mutants of *S. lycopersici*, we exploited recent observations that the protein SGT1 is a key component of signal transduction pathways that lead to disease resistance in plants^{22–24}. VIGS experiments were carried out in which *N. benthamiana* plants were inoculated with a tobacco rattle virus (TRV) vector containing 419 base pairs of the cDNA of *N. benthamiana* SGT1 (TRV:SGT)²⁴. Controls included plants that were inoculated with the TRV vector without insert DNA (TRV:00) and plants that were mock inoculated with 0.01% Tween 80 detergent. After 3 weeks the plants were challenged with either the wild-type *S. lycopersici* strain NEV or the tomatinase-deficient mutant of NEV, 2D2.

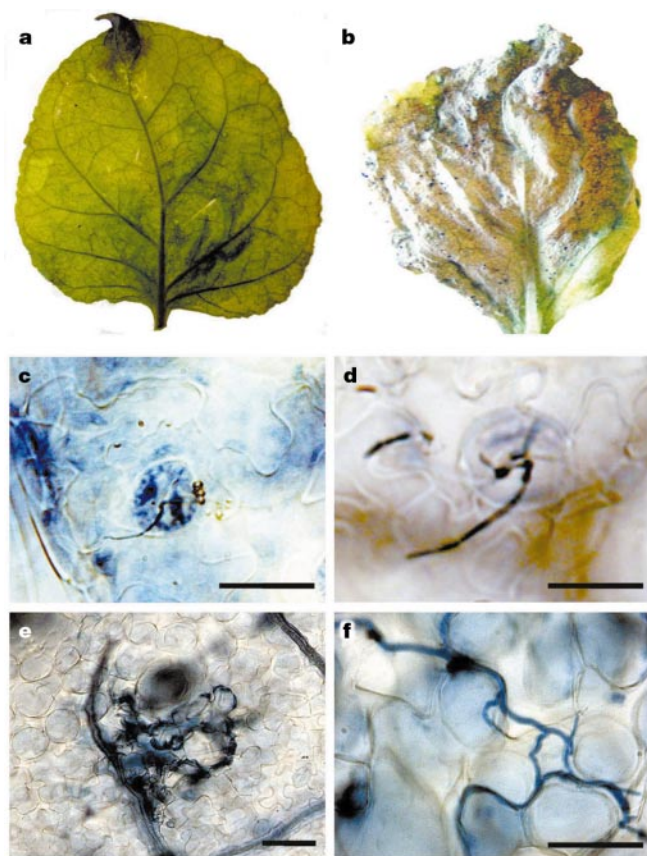


Figure 2 Localized cell death induced by tomatinase-deficient mutants of *S. lycopersici* SGT1-dependent. **a, b**, Appearance of leaves of *N. benthamiana* plants inoculated with TRV:00 (**a**) or TRV:SGT²⁴ (**b**) and then challenged 3 weeks later with 2D2, a tomatinase-deficient mutant of *S. lycopersici* strain NEV. Eight days after fungal inoculation, plants that received the TRV:00 vector showed only mild symptoms, whereas SGT1-silenced plants showed extensive lesions. **c–f**, Leaves stained with trypan blue showing extent of fungal infection in *N. benthamiana* plants inoculated with TRV:00 (**c, e**) or TRV:SGT (**d, f**) 4 d after fungal inoculation. Scale bars, 40 μ m.

TRV:00-inoculated plants that had been subsequently challenged with NEV showed disease symptoms that resembled those in Fig. 1b (data not shown) but were resistant to 2D2 (Fig. 2a, c, e); however, the 2D2 mutant caused necrosis on leaves of SGT1-silenced plants (Fig. 2b) and in early infection was able to enter the stomata and grow around the mesophyll cells without triggering obvious signs of plant cell death (Fig. 2d, f). Thus, the hypersensitive-response-like effect observed when *N. benthamiana* leaves were challenged with 2D2 was dependent on SGT1. These results suggest that the failure of tomatinase-deficient mutants to grow in wild-type *N. benthamiana* leaves is unlikely to be due only to sensitivity to host antibiotics or to nutritional defects, and that tomatinase has another role in establishing disease.

Because colonization of *N. benthamiana* leaves by *S. lycopersici* initially involves intercellular growth¹² and tomatinase is an extracellular enzyme^{10,13–15}, we examined apoplastic fluids recovered from infected leaf tissue for tomatinase activity 6–192 h after inoculation¹². Enzyme activity was readily detectable in apoplastic fluid derived from leaves that had been inoculated with the wild-type *S. lycopersici* strain NEV from 12 h onwards, but not in apoplastic fluid recovered from mock inoculated leaves or from

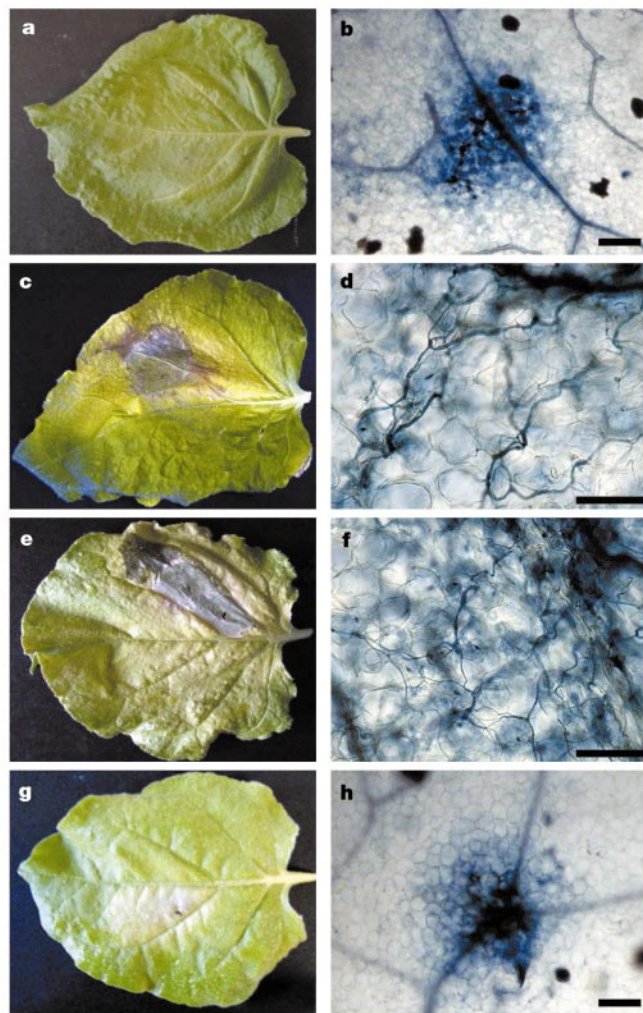


Figure 3 Pre-infiltration with β_2 -tomatine compromises resistance to 2D2. Leaves of *N. benthamiana* were brush-inoculated with 2D2 30 min after pre-infiltration with water (**a, b**), purified tomatinase in water ($25 \mu\text{g ml}^{-1}$; **c, d**), $100 \mu\text{M}$ β_2 -tomatine (**e, f**) or $100 \mu\text{M}$ α -tomatine (**g, h**). Leaves were stained with trypan blue 8 d after inoculation and examined with an Axioskop microscope under bright-field illumination. Scale bars, 40 μ m.

leaves that had been challenged with tomatinase-deficient mutants of NEV (data not shown). Thus, tomatinase is secreted in an active form into the apoplast during infection of *N. benthamiana* leaves by wild-type *S. lycopersici*.

To investigate the role of tomatinase in more detail, leaves of *N. benthamiana* were infiltrated with water, purified tomatinase, β_2 -tomatine or α -tomatine and then inoculated 30 min later with the wild-type *S. lycopersici* strain NEV or the tomatinase-deficient NEV mutant 2D2. Leaves that had been pretreated with purified tomatinase or with β_2 -tomatine were susceptible to the tomatinase-

deficient mutant in the areas that had been pre-infiltrated (Fig. 3c–f), whereas leaves that had received water or α -tomatine were still resistant to the mutant (Fig. 3a, b, g, h). The wild-type *S. lycopersici* strain NEV showed normal disease symptoms on all leaves (data not shown). Control leaves that received the various pre-infiltration treatments but were mock inoculated with 0.01% Tween 80 detergent instead of fungal inoculum appeared healthy, with the exception of α -tomatine-infiltrated leaves, which showed mild chlorosis in the zone of infiltration. We conclude that pre-infiltration with β_2 -tomatine but not α -tomatine permits the growth of 2D2 in *N. benthamiana*. Thus, the effects of tomatinase on disease are mediated by enzymatic hydrolysis products rather than by the protein *per se*. Because these products are likely to be generated in the apoplast, suppression of plant defences must involve signal transduction from the apoplast to the interior of the plant cells.

To determine whether tomatinase and its hydrolysis product β_2 -tomatine interfere with resistance responses induced by pathogens other than *S. lycopersici*, leaves of *N. benthamiana* plants expressing the tomato disease resistance gene *Pto* (ref. 25) were pre-infiltrated with water, purified tomatinase, β_2 -tomatine or α -tomatine. We carried out transient gene expression assays by infiltration with *Agrobacterium tumefaciens* expressing the matching bacterial avirulence gene *avrPto* from *Pseudomonas syringae* pv. *tomato*²⁶. *Pto*-mediated recognition of *avrPto* has been shown to function in transgenic *N. benthamiana* plants²⁷. We observed hypersensitive-response-associated necrosis in leaves that had been treated with water or α -tomatine (Fig. 4c, d). This necrosis was reduced markedly in leaves that had been pre-infiltrated with either tomatinase or β_2 -tomatine (Fig. 4a, b). The pretreatments did not have any effect on transmission of transfer DNA, as assessed by infiltration of *N. benthamiana* leaves with *A. tumefaciens* expressing green fluorescent protein (GFP) under the control of the 35S promoter (ref. 28 and Fig. 4e–i).

We then investigated the effects of the pre-infiltration treatments on growth of *P. syringae* pv. *tabaci* expressing *avrPto* in leaves of wild-type and *Pto*-transgenic *N. benthamiana* plants. Leaves were pretreated with water, purified tomatinase, β_2 -tomatine or α -tomatine as above, and then infiltrated with *P. syringae* pv. *tabaci* (*avrPto*). Bacterial growth in the zone of infiltration was assessed 0, 2 and 4 d after inoculation (Fig. 4j). Bacterial growth in leaves of *N. benthamiana* expressing the *Pto* gene (that is, in the incompatible interaction) was around 10-fold higher 2 and 4 d after inoculation in leaves that had been pretreated with tomatinase or β_2 -tomatine, as compared with leaves that had received water or α -tomatine (Fig. 4j, top). Leaves that had been pretreated with tomatinase or β_2 -tomatine also showed a reduced hypersensitive response. Effects on the compatible interaction were less clear, although bacterial growth seemed to be slightly higher in leaves that had been pretreated with tomatinase or β_2 -tomatine (Fig. 4j, bottom). This difference was reproduced in two more experiments.

Traditionally, enzymes such as tomatinase have been regarded as detoxifying agents that permit fungal growth *in planta* by degrading preformed host antibiotics³. By implication these preformed compounds, which are often sequestered in the outer protective layers of plant tissues, are generally considered to be inert chemical barriers to microbial attack. Our results indicate that tomatinase has a dual function and is required not only for saponin detoxification but also for suppression of fundamental disease resistance processes. *N. benthamiana* is not known to produce α -tomatine and its saponin content has not been characterized. However, other *Nicotiana* species synthesize steroidal glycoalkaloids²⁹ and steroidal saponins³⁰. From our knowledge of the substrate specificity of tomatinase^{10,13–15}, we predict that the endogenous substrate in *N. benthamiana* will be a steroidal glycoalkaloid or glycosylated steroid with a terminal D-glucose molecule. Efforts to identify this substrate are underway. A future challenge will be the biochemical and genetic elucidation of this suppression phenomenon in *N. benthamiana*. A

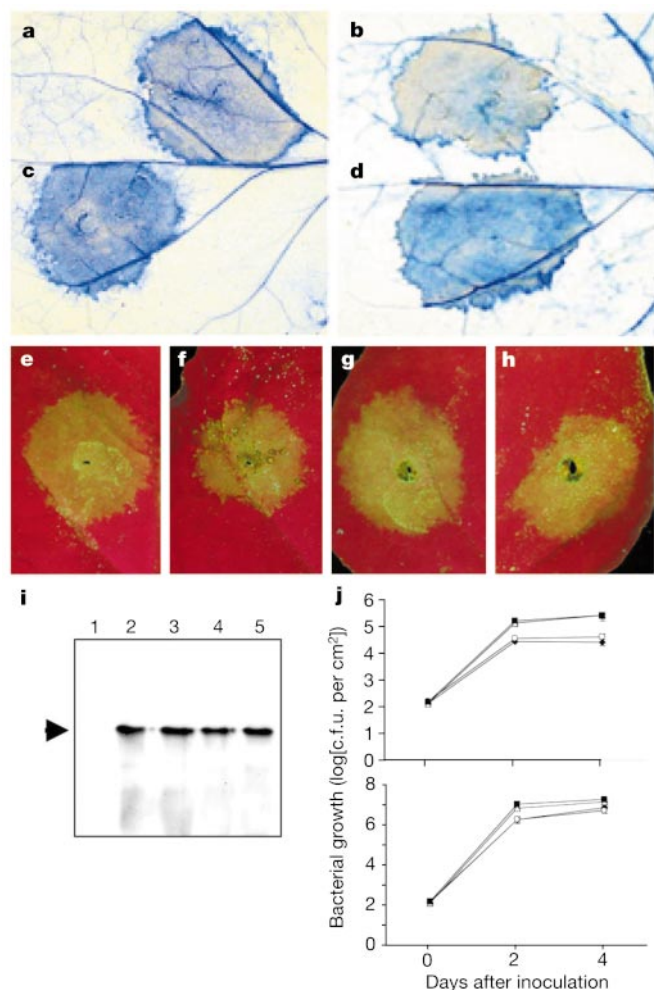


Figure 4 β_2 -Tomatine mediates suppression of bacterial disease resistance in *N. benthamiana*. **a–d**, Hypersensitive-response-associated necrosis detected by trypan blue staining of leaves of *Pto*-expressing *N. benthamiana* transformants. Leaves were pre-infiltrated with purified tomatinase in water (25 μ g ml^{−1}; **a**), 100 μ M β_2 -tomatine (**b**), water (**c**) or 100 μ M α -tomatine (**d**) and then infiltrated with *A. tumefaciens* containing *avrPto* and stained with trypan blue after 4 d. **e–h**, Leaves pretreated as in **a–d** and then infiltrated with *A. tumefaciens* carrying the *GFP* gene. Leaves were viewed under ultraviolet light 4 d after *A. tumefaciens* infiltration. **i**, Western blot analysis of *GFP* expression. Proteins were extracted from six 0.6-cm² leaf discs per treatment, separated by SDS–PAGE and subjected to western blot analysis using anti-*GFP* antibodies. **j**, Growth of *P. syringae* pv. *tabaci* expressing *avrPto* in *Pto*-transgenic (top) and wild-type (bottom) *N. benthamiana* plants. Four-week-old plants were pre-infiltrated with water (open circles), 100 μ M α -tomatine (solid diamonds), purified tomatinase (25 μ g ml^{−1}; solid squares) or 100 μ M β_2 -tomatine (open triangles) and then infiltrated with *P. syringae* pv. *tabaci* (*avrPto*; 10⁴ c.f.u. per ml). Bacteria were recovered from leaf discs and growth determined as c.f.u. per cm² of leaf tissue. Each time point represents the mean for six 0.6-cm² leaf discs taken from each of three plants. Error bars represent standard deviations.

second key area of interest will be establishing whether other pathogens, and possibly also symbionts, use similar mechanisms to colonize their plant hosts. □

Methods

Fungal strains and biochemical procedures

S. lycopersici strains, culture conditions, genomic DNA extraction methods, tomatinase assays and tomatinase purification procedures have been described^{12,15}. We purchased α -tomatine from Sigma and prepared β_2 -tomatine from α -tomatine by enzymatic hydrolysis¹⁵.

Plant infection

All plant inoculations involved a minimum of three leaves from each of three plants, and each experiment was carried out at least three times. *N. benthamiana* was grown under conditions of 8 h light/16 h dark at 22 °C with 70% relative humidity. Leaves of 30-day-old plants were brush-inoculated with either spore suspensions of *S. lycopersici* strains (1×10^6 spores per ml) in 0.01% Tween 80 detergent or detergent only¹². For the infiltration experiments, leaves were pre-infiltrated with either water, purified tomatinase in water (25 $\mu\text{g ml}^{-1}$), β_2 -tomatine or α -tomatine (both at a concentration of 0.1 mM in water, diluted from a 10 mM stock solution in methanol) and then inoculated with *S. lycopersici* after 30 min. For transient gene expression experiments, *N. benthamiana* leaves were infiltrated with *A. tumefaciens* strain C58C1 containing binary constructs with either *AvrPto*²⁶ or *GFP*²⁸. We carried out inoculations with *P. syringae* pv. *tabaci* strains according to published methods. Six 0.6-cm² discs from each leaf were taken at 0, 2 and 4 d after inoculation and bacterial growth was monitored as described²⁴. We carried out VIGS experiments as described²⁴.

RT-PCR

We extracted RNA using TRI reagent (Sigma) and resuspended it in water. Reverse transcriptase reactions were carried out with 2 μg of total RNA for 5 min at 70 °C. We used 1 μl of the reverse transcription reaction as a template in 20- μl PCR reactions, with 5 U of Taq polymerase (Gibco-BRL) per reaction. PCR was run for 24 cycles of 30 s at 94 °C, 30 s at 50 °C, and 1.5 min at 72 °C. We used the following primers: *PR1* sense, 5'-TCGTCCTTTGTAGCTCTGTAGGTG-3', and antisense, 5'-TAGATTCTCGTAATCTCAGTCTCT-3'; *PR5* sense, 5'-ATGCAACAACCCGTGTACTG-3', and antisense, 5'-TTGGTGCAAGTGAAAGTGCG-3'; *AOS* sense, 5'-CTCAAGTCATCTCGAAACCG-3', and antisense, 5'-AGCTTCAACGAGAATCTCACCG-3'; *actin*, sense, 5'-ATGGCAGACGGTGAGGATATTC-3', and antisense, 5'-GCCTTTGCAATCCACATCTGTTG-3'. The RT-PCR products were separated by agarose gel electrophoresis and detected by staining with ethidium bromide. Control PCR reactions (mock reverse-transcribed RNA preparations or lacking the primers) did not generate fragments of the appropriate sizes.

Received 21 May; accepted 13 June 2002; doi:10.1038/nature00950.

- Hammond-Kosack, K. & Jones, J. D. G. in *Biochemistry and Molecular Biology of Plants* (eds Buchanan, B. B., Gruissem, W. & Jones, R. L.) 1102–1156 (American Society of Plant Physiologists, Rockville, Maryland, 2001).
- Flor, H. H. Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**, 275–296 (1971).
- Morrissey, J. P. & Osbourn, A. E. Fungal resistance to plant antibiotics as a mechanism of pathogenesis. *Microbiol. Mol. Biol. Rev.* **63**, 708–724 (1999).
- Doke, N. et al. in *Molecular Genetics of Host-Specific Toxins in Plant Disease* (eds Kohmoto, K. & Yoder, O. C.) 331–341 (Kluwer, Dordrecht, 1998).
- Shiraishi, T. et al. in *Molecular Genetics of Host-Specific Toxins in Plant Disease* (eds Kohmoto, K. & Yoder, O. C.) 343–349 (Kluwer, Dordrecht, 1998).
- Heath, M. C. Hypersensitive response-related death. *Plant Mol. Biol.* **44**, 321–334 (2002).
- Roddick, J. The steroidal glycoalkaloid tomatine. *Phytochemistry* **13**, 9–25 (1974).
- Arneson, P. A. & Durbin, R. D. Studies on the mode of action of tomatine as a fungitoxic agent. *Plant Physiol.* **43**, 683–686 (1968).
- Arneson, P. A. & Durbin, R. D. The sensitivity of fungi to α -tomatine. *Phytopathology* **58**, 536–537 (1968).
- Sandrock, R. W., DellaPenna, D. & VanEtten, H. D. Purification and characterization of β_2 -tomatinase, an enzyme involved in the degradation of α -tomatine and isolation of the gene encoding β_2 -tomatinase from *Septoria lycopersici*. *Mol. Plant Microbe Interact.* **8**, 960–970 (1995).
- Sandrock, R. W. & VanEtten, H. D. Fungal sensitivity to and enzymatic degradation of the phytoanticipin α -tomatine. *Phytopathology* **88**, 137–143 (1998).
- Martin-Hernandez, A. M., Dufresne, M., Hugouvieux, V., Melton, R. & Osbourn, A. Effects of targeted replacement of the tomatinase gene on the interaction of *Septoria lycopersici* with tomato plants. *Mol. Plant Microbe Interact.* **13**, 1301–1311 (2000).
- Arneson, P. A. & Durbin, R. D. Hydrolysis of tomatine by *Septoria lycopersici*: a detoxification mechanism. *Phytopathology* **57**, 1358–1360 (1967).
- Durbin, R. D. & Uchytel, T. F. Purification and properties of a fungal β -glucosidase acting on α -tomatine. *Biochim. Biophys. Acta.* **191**, 176–178 (1969).
- Osbourne, A. E., Bowyer, P., Lunness, P., Clarke, B. & Daniels, M. Fungal pathogens of oat roots and tomato leaves employ closely related enzymes to detoxify different host plant saponins. *Mol. Plant Microbe Interact.* **8**, 971–978 (1995).
- Sandrock, R. W. & VanEtten, H. D. The relevance of tomatinase activity in pathogens of tomato: disruption of the β_2 -tomatinase gene in *Colletotrichum coccodes* and *Septoria lycopersici* and heterologous expression of the *Septoria lycopersici* β_2 -tomatinase in *Nectria haematococca*, a pathogen of tomato fruit. *Physiol. Mol. Plant Pathol.* **58**, 159–171 (2001).
- Baulcombe, D. C. Fast forward genetics based on virus-induced gene silencing. *Curr. Opin. Plant Biol.* **2**, 109–113 (1999).

- Tronchet, M., Ranty, B., Marco, Y. & Roby, D. HSR203 antisense suppression in tobacco accelerates development of hypersensitive cell death. *Plant J.* **27**, 115–127 (2001).
- Yoshioka, K. et al. Environmentally sensitive, SA-dependent defense responses in the cpr22 mutant of *Arabidopsis*. *Plant J.* **26**, 447–459 (2001).
- Froehlich, J. E., Itoh, A. & Howe, G. A. Tomato allene oxide synthase and fatty acid hydroperoxide lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes of chloroplast envelope. *Plant Physiol.* **125**, 306–317 (2001).
- Ziegler, J., Keinänen, M. & Baldwin, I. T. Herbivore-induced allene oxide synthase transcripts and jasmonic acid in *Nicotiana attenuata*. *Phytochemistry* **58**, 729–738 (2001).
- Austin, M. J. et al. Regulatory role of SGT1 in early R gene-mediated plant defenses. *Science* **292**, 2077–2080 (2002).
- Azevedo, C. et al. The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance. *Science* **292**, 2073–2076 (2002).
- Peart, J. et al. *SGT1* is required for host and nonhost disease resistance in plants. *Proc. Natl. Acad. Sci. USA* advance online publication, 15 July 2002 (DOI 10.1073/pnas.152330599).
- Martin, G. B. et al. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**, 1432–1436 (1993).
- Scofield, S. R. et al. Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science* **274**, 2063–2065 (1996).
- Rommens, C. M. T., Salmeron, J. M., Oldroyd, G. E. D. & Staskawicz, B. J. Intergenic transfer and functional expression of the tomato disease resistance gene *Pto*. *Plant Cell* **7**, 1537–1544 (1995).
- Voinnet, O. & Baulcombe, D. Systemic signalling in gene silencing. *Nature* **389**, 553–553 (1997).
- Singh, S., Khanna, N. M. & Dhar, M. M. Solaplumbin, a new anticancer glycoside from *Nicotiana plumbaginifolia*. *Phytochemistry* **13**, 2020–2022 (1974).
- Grünweller, S., Schröder, E. & Kesselmeier, J. Biological activities of furostanol saponins from *Nicotiana tabacum*. *Phytochemistry* **29**, 2485–2490 (1990).

Acknowledgements

We thank R. Sandrock and H. Van Etten for strains of *S. lycopersici*; P. Moffett for supplying *SGT1*-silenced *N. benthamiana* plants; and D. Holden and J. Rathjen for criticism of the manuscript. K.B. is supported by a Marie Curie European Community fellowship and the Sainsbury Laboratory is supported by the Gatsby Charitable Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.O. (e-mail: annie.osbourn@sainsbury-laboratory.ac.uk).

Patched acts catalytically to suppress the activity of Smoothened

J. Taipale, M. K. Cooper, T. Maiti & P. A. Beachy

Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

Mutations affecting the transmembrane proteins Patched (Ptc) or Smoothened (Smo) that trigger ligand-independent activity of the Hedgehog (Hh) signalling pathway are associated with human tumours such as basal cell carcinoma (BCC) and medulloblastoma^{1–3}. Despite extensive genetic studies demonstrating the importance of these receptor components in embryonic patterning and cancer, the mechanism by which Ptc regulates Smo is not understood⁴. Here we report that Ptc and Smo are not significantly associated within Hh-responsive cells. Furthermore, we show that free Ptc (unbound by Hh) acts sub-stoichiometrically to suppress Smo activity and thus is critical in specifying the level of pathway activity. Patched is a twelve-transmembrane protein with homology to bacterial proton-driven transmembrane molecular transporters; we demonstrate that the function of Ptc is impaired by alterations of residues that are conserved in and required for function of these bacterial transporters. These results suggest that the Ptc tumour suppressor functions normally as a transmembrane molecular transporter, which acts

indirectly to inhibit Smo activity, possibly through changes in distribution or concentration of a small molecule.

Activation of the Hh pathway is triggered by binding of the Hh protein to Ptc^{5,6}. This interaction is equivalent to genetic loss of Ptc function, as loss of Ptc also results in pathway activation, even in the absence of Hh signal^{7,8}. Whether resulting from loss of Ptc or the presence of Hh, however, pathway activation requires the activity of the seven-transmembrane protein Smo⁹. Hh signal thus activates the Hh pathway by suppressing the negative regulatory action of Ptc on Smo, and Smo then activates Hh target genes through the Ci/GLI family of transcription factors (for reviews, see refs 1–4, 9). The actual mechanism of suppression of Smo activity by Ptc remains unclear. One possibility is that Ptc suppresses Smo activity by direct interaction, and that Hh relieves this suppression by binding to Ptc without causing dissociation of the Smo–Ptc complex⁵ (see Fig. 1a). This model postulates a positive influence of Hh-bound Ptc on Smo, a feature that is not required from genetic evidence as pathway activation can occur in the absence of Ptc^{7,8}. A simpler stoichiometric model would involve a dissociating heteromeric receptor, which would dissociate on Hh-binding and pathway activation, thus releasing Ptc suppression of Smo (Fig. 1b). The principal features of all stoichiometric models of this pathway are that regulation of Smo activity requires at least equal levels of Ptc, and that Ptc must associate with Smo in cells. The alternative to these

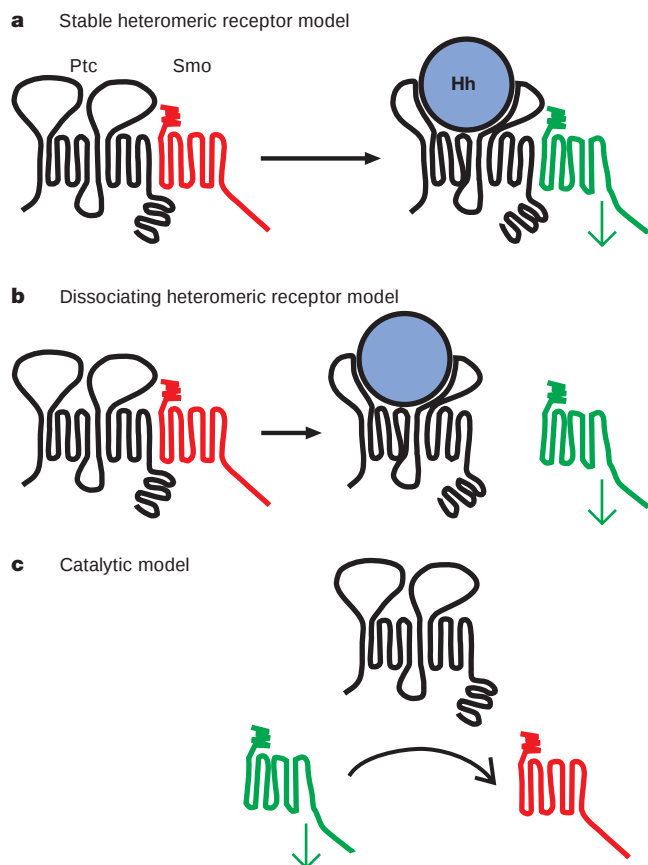


Figure 1 Models of Hh receptor function. **a, b**, Stoichiometric models. In a stable heteromeric receptor (**a**), Hh-binding to Ptc induces a conformational change in a stable Ptc–Smo complex, allowing Smo activity⁵. In a dissociating heteromeric receptor (**b**) Ptc and Smo are physically associated in the unstimulated state (left), but stimulation by Hh induces dissociation, generating free, active Smo molecules (right). **c**, Catalytic model. Free Ptc molecules catalyse the conversion of active Smo to an inactive form. Binding of Hh to Ptc inhibits this catalytic activity.

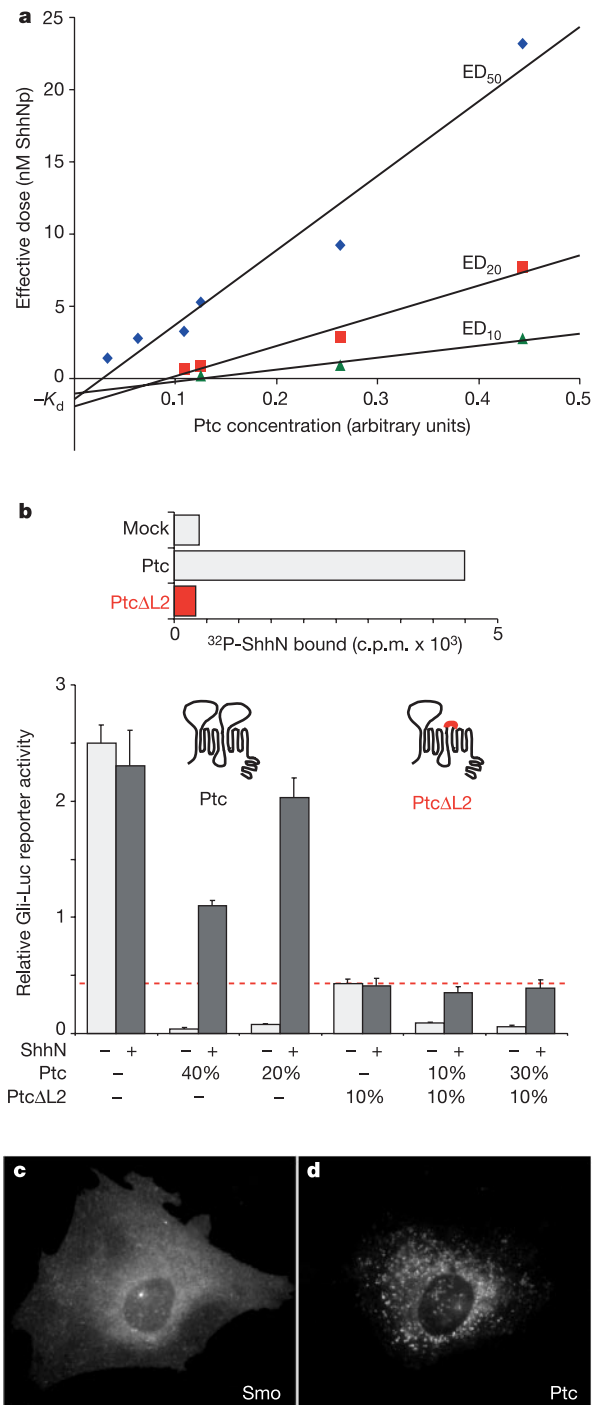


Figure 2 Free Ptc governs Hh pathway activity. **a**, Increasing Ptc levels decrease ShhNp sensitivity. PZA^{Ptc-/-} cells were transfected with Ptc–Renilla constructs and reporters for Hh pathway activity and transfection efficiency, and were treated with ShhNp. ShhNp concentrations required for 10% (ED_{10}), 20% (ED_{20}) or 50% (ED_{50}) of maximal pathway activity are plotted as a function of the concentration of Ptc expressed (normalized Renilla luciferase activity). Note the linear relationship (see Supplementary Information) with y-intercepts yielding ShhNp–Ptc K_d (1.5 ± 0.4 nM; ~ 2 nM reported¹²). **b**, Ptc Δ L2 does not bind Shh but dominantly blocks Shh response. PZA^{Ptc-/-} cells were transfected to express reporters, and Ptc Δ L2, Ptc, and ShhN as indicated. Note that Ptc Δ L2 establishes an upper limit of pathway activity (red line) that cannot be exceeded by co-expression of Ptc and high levels of ShhN. Error bars indicate one standard deviation (quadruplicate wells). 32 P-labelled ShhN binds to COS1 cells expressing Ptc but not Ptc Δ L2 (top). **c, d**, Representative images of C3H10T1/2 cells 48 h after transfection with Smo–yellow fluorescent protein or Ptc–green fluorescent protein reveal that Smo is localized largely to the plasma membrane, whereas Ptc is found in intracellular vesicles.

stoichiometric models is that Ptc could act catalytically to regulate Smo, either directly or indirectly (Fig. 1c). The principal features of a catalytic model would be that sub-stoichiometric levels of Ptc may suffice for regulation of Smo, and that pathway activation would depend on reduction of catalytically active Ptc (free Ptc; that is, unbound by Hh) to a level below the threshold that is permissive of Smo activation. All of these models are consistent with genetic evidence. Some experiments have favoured the heteromeric receptor⁵; others have demonstrated that Smo levels can be significantly elevated in *Drosophila* tissues without activation of the Hh response^{10,11}, suggesting either a cellular reservoir of excess Ptc or a sub-stoichiometric mode of Ptc action. Because the mechanism of Ptc inhibition of Smo activity remains unresolved^{4,9}, we set out to test explicitly these models in cells responsive to Hh.

A strong prediction of the stable heteromeric receptor model (Fig. 1a) is that, in cells where Smo is adequately regulated by Ptc, additional Ptc protein should have no impact on pathway activity if fractional occupancy of Ptc by Hh remains constant. This is because the number of stable heteromeric receptors would be limited by the number of Smo molecules, and additional Ptc protein not found in complex with Smo (free Ptc) should be irrelevant. To test this prediction we expressed increasing amounts of Ptc in PZA^{Ptc^{-/-}} (see Methods) cells, which lack endogenous Ptc, and monitored pathway activity at various Ptc levels (see Methods) as a function of the concentration of purified Sonic hedgehog amino-terminal domain (ShhNp). We found that in PZA(Ptc^{-/-}) cells, an increase

in Ptc expression at a constant ShhNp concentration depressed pathway activity, and that corresponding increases in ShhNp concentration were required to maintain the pathway at a constant 10%, 20% or 50% of maximal activity; at the highest concentrations of Ptc tested, the half-maximal effective dose (ED₅₀) for ShhNp was more than tenfold above the reported dissociation constant (*K_d*) (2 nM; ref. 12) of the ShhNp–Ptc interaction (Fig. 2a). Thus, contrary to expectation from the stable heteromeric receptor model, additional Ptc protein exceeding the level required for Smo regulation is not irrelevant, and instead seems to critically affect pathway activation.

To test directly the importance of free Ptc in regulation of Smo, we expressed Ptc protein lacking a segment of the extracellular loop between transmembrane spans 7 and 8 (PtcΔL2; Fig. 2b). As previously reported¹³ we found that this protein is normally expressed and retains the ability to suppress the pathway, albeit with reduced specific activity. In contrast to wild type, however, PtcΔL2 is incapable of ligand binding and is not responsive to stimulation by Shh (added ShhNp or co-transfected ShhN). In cells simultaneously expressing varying levels of Ptc and PtcΔL2, the stable heteromeric receptor model would predict that an increase in relative levels of wild-type Ptc should displace PtcΔL2 from heteromeric receptor complexes, and that the overall responsiveness to Shh therefore should increase. We found, however, that increased expression of wild-type Ptc protein could not increase the level of pathway activity beyond the level specified by PtcΔL2 alone, even under conditions of maximal Shh stimulation (Fig. 2b, red line). These results indicate that the suppressing influence of Shh-insensitive Ptc cannot be diluted out by competition from wild-type Ptc, contrary to expectation from the stable heteromeric receptor model. The lack of Hh stimulation in the presence of PtcΔL2 also suggests that Hh contributes to pathway activation solely through its suppression of Ptc activity. That is, consistent with genetic evidence⁷, neither Hh nor the Hh–Ptc complex are required to positively activate Smo.

Although the ability of free Ptc to regulate Smo activity (Fig. 2a, b) directly contradicts the stable heteromeric receptor model (Fig. 1a), it does not rule out a heteromeric receptor that dissociates on binding to Hh (Fig. 1b). However, contrary to expectation from this model, we failed to detect specific association between Smo and Ptc in unstimulated PZA^{Ptc^{-/-}} cells by immunoprecipitation even after crosslinking with membrane-soluble crosslinking agent (see Supplementary Information). Furthermore, we found that in mammalian cells, as reported by others in *Drosophila*^{11,14}, Ptc and Smo localizations are quite distinct, with Smo prominently localized to the plasma membrane (Fig. 2c), and Ptc mostly in intracellular vesicular structures resembling endosomes (Fig. 2d), with barely detectable amounts found in the plasma membrane.

To distinguish further stoichiometric from non-stoichiometric modes of Ptc action, we examined the levels of Ptc required to regulate Smo using Ptc and Smo proteins identically tagged with a triple-Myc epitope and co-expressed in PZA^{Ptc^{-/-}} cells (Fig. 3a). Detection of Ptc and Smo proteins, expressed at near-physiological levels (see Methods), required bulk immunoprecipitation followed by highly sensitive immunoblotting. Ptc expression at levels that constituted only a small fraction of Smo protein levels was nevertheless sufficient to substantially reduce the activity of Smo protein. Thus, for example, a 1:45 ratio of Ptc to Smo expression constructs resulted in a nearly 80% reduction of Smo activity, despite a great excess of Smo protein (lane 5, Fig. 3a).

For a more quantitative determination of Ptc:Smo regulatory stoichiometry we expressed increasing levels of Renilla luciferase-tagged Smo in PZA^{Ptc^{-/-}} cells, and for each concentration of Smo we determined the concentration of co-transfected Ptc–Renilla that resulted in 50% of maximal pathway activity (Fig. 3b). A plot of Smo levels as a function of Ptc levels required for 50% pathway suppression displayed a linear relationship with a slope of approxi-

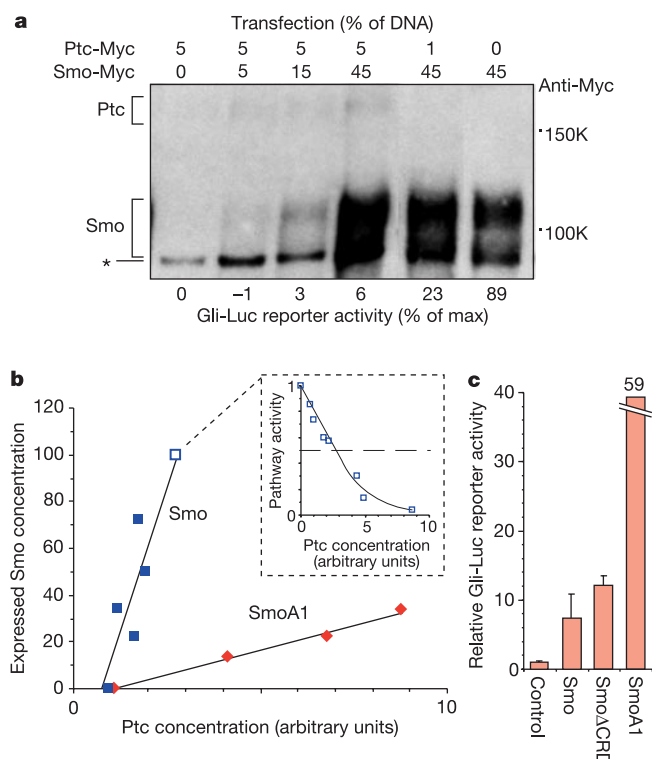


Figure 3 Ptc acts sub-stoichiometrically to regulate Smo. **a**, PZA^{Ptc^{-/-}} cells were transfected with reporters and triple-Myc-tagged Smo and/or Ptc expression constructs as indicated. Reporter assays were normalized to give percentage of maximal Hh pathway activity, as shown below the anti-Myc immunoblot. Asterisk indicates a background band slightly below the faster-migrating form of Smo. *M_r*, relative molecular mass, in thousands (K). **b**, Ptc–Renilla fusion protein levels required to suppress Gli-Luc reporter activity by 50% (Ptc_{50%}) in PZA^{Ptc^{-/-}} cells determined for several concentrations of Smo–Renilla (squares) and SmoA1–Renilla (diamonds). Inset Ptc_{50%} determination at one Smo concentration (open square). **c**, SmoΔCRD activity is similar to that of wild-type Smo. The reporter assay was performed in NIH-3T3 cells that express endogenous Ptc; partially Ptc-resistant oncogenic SmoA1 is shown for comparison.

mately 50 (Fig. 3b), indicating that half-maximal pathway activity is observed when Smo is in 50-fold molar excess of Ptc. Similar analysis with oncogenically activated SmoA1 also revealed a linear relationship with a slope of about 4.2, consistent with the decreased sensitivity of this protein to Ptc inhibition. These linear relationships conform fully to the quantitative predictions of a catalytic model (see Supplementary Information). We also considered the possibility that Ptc might regulate Smo activity through the N-

terminal cysteine-rich domain (CRD) of Smo, analogous to the way that Wnt proteins regulate the Smo-related Frizzled proteins by binding to Frizzled CRDs¹⁵. However, deletion or alteration of the CRD did not produce a significant difference in Smo activity or in sensitivity to suppression by Ptc (Fig. 3c).

The ability of Ptc to act sub-stoichiometrically in suppressing Smo activity strongly supports a catalytic mode of Ptc action, but does not identify the biochemical mechanism of this action. Several recent studies in *Drosophila* have noted a signalling-related increase in Smo phosphorylation and a post-transcriptional increase in levels of Smo protein^{10,11,16}. However, an experimentally induced increase in Smo protein to a level higher than that found in cells where the pathway is fully activated does not result in pathway activation (see Fig. 2 of ref. 11), indicating that another mechanism must account for Hh-induced increase in Smo activity. We did not observe substantial regulation of mouse Smo-Renilla or Smo-Myc protein levels in mammalian cells by Ptc or Shh (Fig. 3a; not shown), and mutations in potential conserved phosphorylation sites in mouse or *Drosophila* Smo had limited or no effect on signalling activity (J.T., L. Lum, D. von Kessler and P.A.B., unpublished observation), suggesting that neither changes in stability nor phosphorylation of Smo are necessary for primary activation of the pathway. Alternative suggestions as to the nature of the catalytic activity of Ptc emerge from similarity between Ptc and the resistance, nodulation, division (RND) family of bacterial proton gradient-driven transmembrane molecular transporters. In bacteria, the RND proteins function in the export of nodulation and quorum-sensing factors and as multidrug efflux pumps, removing antibiotics, toxic organic compounds, and metal ions from cells¹⁷. The RND transporters share with Ptc a similar domain structure of 12 predicted transmembrane spans (TMs), with large extracellular loops protruding between the first and second, and seventh and eighth transmembrane spans¹⁷. In addition, a protein more closely related to Ptc, Niemann-Pick C1 disease protein (NP-C1), has been shown to function as a transmembrane molecular transporter¹⁸. Another related protein, Dispatched¹⁹, is required for long-range Hh signalling and seems to act in release of Hh protein from the cell, perhaps by extrusion of the palmitate and cholesterol adducts of Hh from the membrane. Of particular interest in the RND family, residues GxxxD are highly conserved at a position in the middle of TM4, and alteration of the aspartate reduces transporter function, although some channel activity remains^{20,21}. Ptc proteins also contain these conserved residues. Notably, of the six missense mutations in Ptc that are known to be associated with Gorlin's syndrome^{22–24}, three affect these two conserved residues (Fig. 4a), strongly suggesting the existence of a function that is conserved between the RND family and Ptc. Functional testing in PZA^{Ptc-/-} cells of altered forms of Ptc proteins incorporating mutations in these residues demonstrates that these proteins are 7–8-fold less active than wild-type Ptc in suppression of pathway activity (Fig. 4b).

We have shown that Ptc does not extensively associate with Smo and, in several distinct experiments, that levels of free Ptc protein determine the degree of pathway activity as well as the amount of Hh ligand required for pathway stimulation. We have also shown that sub-stoichiometric levels of Ptc suffice to regulate Smo by a mechanism that may be closely related to the action of bacterial RND transporters. Several lines of evidence indicate that Smo exists in a balance between active and inactive states^{11,16,25}. These activity states are probably conformationally distinct²⁵, analogous to the activity states of G-protein-coupled receptors that are activated by proteins or by small-molecule ligands. In general, such receptors either use large N-terminal domains to bind protein ligands, or may be activated by small molecules, in which case these N-terminal domains are generally absent. Although Smo has a large N-terminal domain, this domain is dispensable for function (Fig. 3c), suggesting that small molecules might gate Smo activity. Interest-

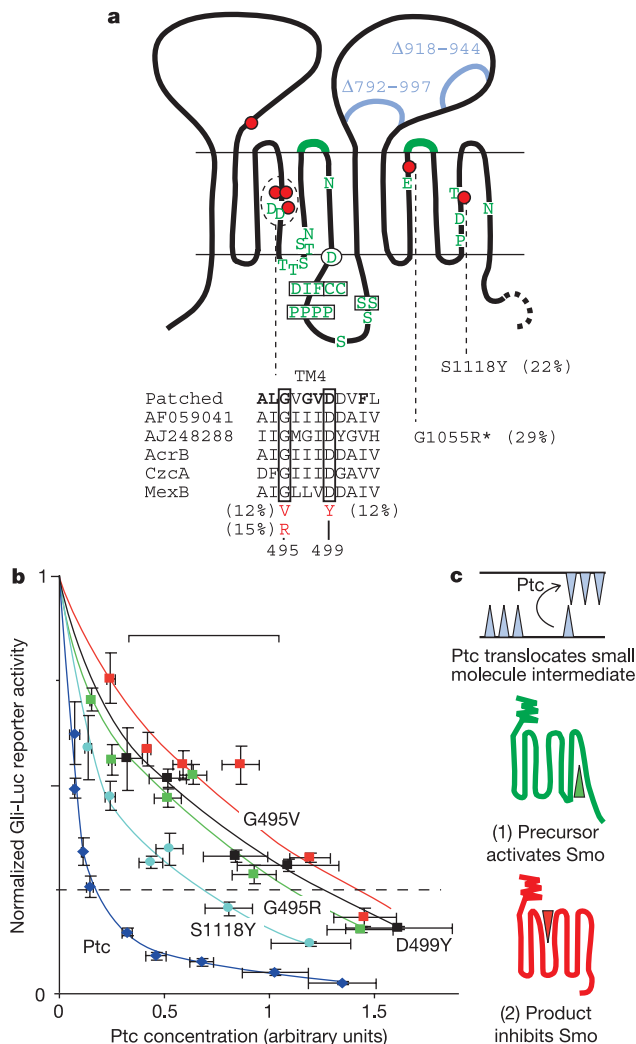


Figure 4 Ptc may function as a transmembrane molecular transporter. **a**, Gorlin's syndrome missense mutations^{22–24} (red dots) map to conserved residues in TM4 (see boxes in alignment) that are required for molecular transporter function in the RND family^{20,21}. Specific activities of Gorlin's syndrome mutant proteins in pathway suppression are given as a percentage of wild-type Ptc (see **b**). The G1055R mutation (asterisk) also resulted in poor (<25%) protein expression. Other mutations in conserved Ptc residues (bold) including the 'sterol sensor' mutation (circled; see Supplementary Information), that did not significantly affect activity are in green. Blue lines indicate Ptc deletion mutants that suppressed the pathway but were unresponsive to Shh. **b**, Physiological levels (horizontal bracket) of Gorlin's syndrome mutant Ptc proteins fail to regulate adequately the Hh pathway in PZA^{Ptc-/-} cells. Hh pathway activity is shown as a function of indicated Ptc Renilla-fusion protein concentrations. **c**, Model of Hh receptor function. In the absence of Hh, Ptc translocates a small molecule, resulting in loss of Smo activator (1) or gain of Smo inhibitor (2). The small fraction of Ptc found co-localized with Smo (see text) could suffice for Smo regulation, or the small molecule intermediate could be transported. The intermediate might also act more indirectly, for example, by affecting vesicular trafficking events that may impact on the activity cycle of Smo^{28–30}.

ingly, several small molecules that act as Hh-pathway agonists or antagonists (cyclopamine) have been identified, and these compounds seem to act by binding directly to Smo and modulating its activity (ref. 25 and J. K. Chen and P.A.B., unpublished observations). In addition, all described activated forms of the Smo protein are resistant both to cyclopamine and to Ptc²⁵, suggesting that Ptc and cyclopamine may act in a similar way to inhibit Smo activity. The simplest model of Hh-receptor function consistent with our data is that in the absence of Hh, Ptc translocates across the membrane bilayer a small molecule that directly or indirectly regulates the Smo activity state, with the small molecule substrate acting either as a Smo agonist before or as an antagonist after translocation by Ptc (Fig. 4c). Such a mechanism of receptor action raises the possibility of therapeutic manipulation of the Hh pathway with the use of physiological agonists or antagonists of Smo, or by small molecules such as cyclopamine that may mimic the action of such small molecules. □

Methods

Cell culture and reagents

The PZA^{Ptc^{-/-}} cell line is a soft-agar cloned tumorigenic derivative of P2^{Ptc^{-/-}} cells²⁵ that has better growth characteristics and higher transfection efficiency than the parental cells. All vertebrate cell lines were cultured in DMEM with 10% calf serum (NIH-3T3) or fetal calf serum (all other cells). ³²P-labelled ShhN binding assay was performed as described⁶. Despite palmitate and cholesterol modifications, the purified Shh N-terminal domain²⁵ (ShhNp) functioned like a soluble factor in signalling assays (see also ref. 26)—the concentration rather than the absolute amount of ShhNp per cell determined pathway activity. The ShhNp protein (estimated at about 10⁸ molecules per cell at 1 nM) was in great excess of Ptc (see below), satisfying the requirement for a constant fractional occupancy of Ptc. In addition, bioassay of ShhNp in culture supernatants after the experiments in Fig. 2a indicated that pathway-stimulating activities were similar regardless of Ptc expression levels, indicating that significant amounts of ShhNp were not sequestered by transfected Ptc (not shown).

Constructs and transfections

A Shh-sensitive firefly luciferase reporter (Gli-Luc reporter) was constructed by cloning eight Gli-binding sites²⁷ to pGL3 (Promega). Two restriction sites were introduced into the carboxy termini of Ptc and Smo, into which Renilla luciferase (Promega), fluorescent proteins (Clontech) and triple-Myc tags were cloned. All fusion and tagged constructs exhibited activities indistinguishable from the corresponding wild-type constructs in PZA^{Ptc^{-/-}} and NIH-3T3 cells. Point mutations, deletions, and Ptc/NP-CI chimaeric sequences were introduced by polymerase chain reaction (PCR) (Pfu polymerase) to the Ptc–Renilla construct. Ptc mutants not defined in Fig. 5a are: D499A, D500A, T511A, T523A, S533A, T537A, S538A, N541A, 557RAFSLQAA/HTFSLFAG, N570A, D585N, 595DIF/AAA, 598CC-AA, 627PPP/AAAA, S669A, S699A, 704SS/AA, E1052A, 1060IGIK/WGIS, P1111A, D1114A, T1119A, N1152A. The CRD of Smo (residues 68–182) was deleted using the structure of Frizzled CRD¹⁵ to predict disulphide bond pattern. Cells were split 1:6 and transfections were carried out on the following day using Fugene6 (0.3 µl per 100 ng DNA) with 0–100% (w/w) effector and control vector (pEGFP-C1), and where indicated 50% (w/w) Gli-Luc and Renilla luciferase (pRL-SV40) or secreted alkaline phosphatase (pSEAP-CT) transfection control vector (20:1 ratio). After transfection in full medium for 2–3 d, the cells were changed to low serum medium (DMEM with 0.5% calf serum, 5 mM HEPES, pH 7.4) with or without ShhNp, and alkaline phosphatase, Renilla luciferase and firefly luciferase activities were determined (Phospha-Light, Tropix; Dual-Luc, Promega) after 2–3 d of additional culture. In all cases, both reporters for pathway activity and protein level were normalized to transfection efficiency. Effective concentrations of Ptc or ShhNp required for a given level of pathway activity were determined by linear interpolation. In Fig. 3b, pathway activity and concentrations of Ptc (Renilla luciferase activity measured for each w/w ratio of Ptc–Renilla construct) and exogenous Smo (measured Renilla-luciferase activity with Ptc–Renilla-derived activity subtracted) were determined 5 d after transfection. Physiological expression range of Ptc was estimated as the level of Ptc capable of >90% suppression of the pathway yet resulting in near-physiological ED₅₀ for ShhNp (see Figs 2a and 4b). Note that the levels of recombinant protein produced by the pRK5 vector in PZA^{Ptc^{-/-}} cells are much lower than those in COS1 cells, where this vector replicates episomally (for example, ~4,000 Ptc molecules per transfected cell as compared with ~200,000 in COS1 cells; estimates derived by determination of transfection efficiency and the number of Ptc-luciferase molecules in extracts, using purified luciferase protein as a standard).

Antibodies

Polyclonal (A14) and monoclonal (9E10) antibodies to the Myc-tag were from Santa Cruz Biotechnology. In Fig. 3a, Myc-tagged protein levels were assayed 2 d after transfection by immunoprecipitation with 9E10 followed by immunoblotting with A14. Pathway activity was determined after 2 d further culture in low serum medium (duplicate plate).

Homology searches

An NCBI conserved domain search (no filter) identified homology between human Ptc

amino acids 879–1025 and the AcrB/AcrD/AcrF bacterial RND transporter protein family (pfam00873). Multiple proteins within this family also have homology to the Ptc domain (pfam02460). Sequences other than MexB and Czca (shown in Fig. 4a) were identified from a pattern-hit-initiated BLAST search using the following pattern derived from mouse Ptc sequence: 3*[FLIVWMAG]-G-3*[FLIVWMAG]-D-X-3*[FLIVWMAG], where 3*[FLIVWMAG] indicates three consecutive Phe, Leu, Ile, Val, Trp, Met, Ala or Gly residues.

Received 28 February; accepted 26 June 2002; doi:10.1038/nature00989.

- Bale, A. E. & Yu, K. P. The hedgehog pathway and basal cell carcinomas. *Hum. Mol. Genet.* **10**, 757–762 (2001).
- Wechsler-Reya, R. & Scott, M. P. The developmental biology of brain tumors. *Annu. Rev. Neurosci.* **24**, 385–428 (2001).
- Taipale, J. & Beachy, P. A. The Hedgehog and Wnt signalling pathways in cancer. *Nature* **411**, 349–354 (2001).
- Kalderon, D. Transducing the Hedgehog signal. *Cell* **103**, 371–374 (2000).
- Stone, D. M. *et al.* The tumour-suppressor gene *patched* encodes a candidate receptor for Sonic hedgehog. *Nature* **384**, 129–134 (1996).
- Fuse, N. *et al.* Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for patched. *Proc. Natl Acad. Sci. USA* **96**, 10992–10999 (1999).
- Ingham, P. W., Taylor, A. M. & Nakano, Y. Role of *Drosophila* patched gene in positional signaling. *Nature* **353**, 184–187 (1991).
- Goodrich, L. V., Milenkovic, L., Higgins, K. M. & Scott, M. P. Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* **277**, 1109–1113 (1997).
- Ingham, P. W. & McMahon, A. P. Hedgehog signalling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059–3087 (2001).
- Ingham, P. W. *et al.* Patched represses the Hedgehog signalling pathway by promoting modification of the Smoothened protein. *Curr. Biol.* **10**, 1315–1318 (2000).
- Denef, N., Neubuser, D., Perez, L. & Cohen, S. M. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* **102**, 521–531 (2000).
- Williams, K. P. *et al.* Functional antagonists of sonic hedgehog reveal the importance of the N terminus for activity. *J. Cell. Sci.* **112**, 4405–4414 (1999).
- Briscoe, J., Chen, Y., Jessell, T. M. & Struhl, G. A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. *Mol. Cell* **7**, 1279–1291 (2001).
- Capdevila, J., Pariente, F., Sampedro, J., Alonso, J. L. & Guerrero, I. Subcellular localization of the segment polarity protein patched suggests an interaction with the wingless reception complex in *Drosophila* embryos. *Development* **120**, 987–998 (1994).
- Dann, C. E. *et al.* Insights into Wnt binding and signalling from the structures of two frizzled cysteine-rich domains. *Nature* **412**, 86–90 (2001).
- Alcedo, J., Zou, Y. & Noll, M. Posttranscriptional regulation of Smoothened is part of a self-correcting mechanism in the Hedgehog signaling system. *Mol. Cell* **6**, 457–465 (2000).
- Tseng, T. T. *et al.* The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* **1**, 107–125 (1999).
- Davies, J. P., Chen, F. W. & Ioannou, Y. A. Transmembrane molecular pump activity of niemann-pick C1 protein. *Science* **290**, 2295–2298 (2000).
- Burke, R. E. *et al.* Dispatched, a novel sterol sensing domain protein dedicated to the release of cholesterol-modified Hedgehog from signaling cells. *Cell* **99**, 803–815 (1999).
- Goldberg, M., Pribyl, T., Juhnke, S. & Nies, D. S. Energetics and topology of Czca, a cation/proton antiporter of the resistance-nodulation-cell division protein family. *J. Biol. Chem.* **274**, 26065–26070 (1999).
- Guan, L. & Nakae, T. Identification of essential charged residues in transmembrane segments of the multidrug transporter MexB of *Pseudomonas aeruginosa*. *J. Bacteriol.* **183**, 1734–1739 (2001).
- Chidambaram, A. *et al.* Mutations in the human homologue of the *Drosophila* Patched gene in caucasian and African-American nevoid basal cell carcinoma syndrome patients. *Cancer Res.* **56**, 4599–4601 (1996).
- Wicking, C. *et al.* Most germ-line mutations in the nevoid basal cell carcinoma syndrome lead to premature termination of the Patched protein, and no genotype–phenotype correlations are evident. *Am. J. Hum. Genet.* **60**, 21–26 (1997).
- Aszterbaum, M. *et al.* Identification of mutations in the human Patched gene in sporadic basal cell carcinomas and in patients with the basal cell nevus syndrome. *J. Invest. Dermatol.* **110**, 885–888 (1998).
- Taipale, J. *et al.* Effects of oncogenic mutations in Smoothened and Patched can be reversed by cyclopamine. *Nature* **406**, 1005–1009 (2000).
- Zeng, X. *et al.* A freely diffusible form of Sonic hedgehog mediates long range signalling. *Nature* **411**, 716–720 (2001).
- Sasaki, H., Hui, C.-C., Nakafuku, M. & Kondoh, H. A binding site for Gli proteins is essential for HNF3β floor plate enhancer activity in transgenics and can respond to Shh *in vitro*. *Development* **124**, 1313–1322 (1997).
- Strutt, H. *et al.* Mutations in the sterol-sensing domain of Patched suggest a role for vesicular trafficking in Smoothened regulation. *Curr. Biol.* **11**, 608–613 (2001).
- Martin, V., Carrillo, G., Torroja, C. & Guerrero, I. The sterol-sensing domain of Patched protein seems to control Smoothened activity through Patched vesicular trafficking. *Curr. Biol.* **11**, 601–607 (2001).
- Ko, D. C., Gordon, M. D., Jin, J. Y. & Scott, M. P. Dynamic movements of organelles containing Niemann-Pick C1 protein: NPC1 involvement in late endocytic events. *Mol. Biol. Cell* **12**, 601–614 (2001).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

We thank J. K. Chen, D. Ginty, D. Leahy, J. Nathans and G. Seydoux for critical review of

the manuscript, and K. Young and R. Gong for technical assistance. P.A.B. is an investigator of the Howard Hughes Medical Institute. This work was supported by an NIH grant.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.A.B. (e-mail: pbeachy@jhmi.edu).

Dynamics of ATP-dependent chromatin assembly by ACF

Dmitry V. Fyodorov & James T. Kadonaga

Section of Molecular Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0347, USA

The assembly of DNA into chromatin is a critical step in the replication and repair of the eukaryotic genome^{1–8}. It has been known for nearly 20 years that chromatin assembly is an ATP-dependent process⁹. ATP-dependent chromatin-assembly factor (ACF) uses the energy of ATP hydrolysis for the deposition of histones into periodic nucleosome arrays, and the ISWI subunit of ACF is an ATPase that is related to helicases^{10,11}. Here we show that ACF becomes committed to the DNA template upon initiation of chromatin assembly. We also observed that ACF assembles nucleosomes in localized arrays, rather than randomly distributing them. By using a purified ACF-dependent system for chromatin assembly, we found that ACF hydrolyses about 2–4 molecules of ATP per base pair in the assembly of nucleosomes. This level of ATP hydrolysis is similar to that used by DNA helicases for the unwinding of DNA¹². These results suggest that a tracking mechanism exists in which ACF assembles chromatin as an ATP-driven DNA-translocating motor. Moreover, this proposed mechanism for ACF may be relevant to the function of other chromatin-remodelling factors that contain ISWI subunits.

A family of chromatin-remodelling factors use the energy of ATP hydrolysis to facilitate nuclear processes in the milieu of chromatin^{1–8}. These factors contain an ATPase subunit that is related to the yeast Swi2/Snf2 protein, which is a member of the DEAD/H superfamily of DNA and RNA helicases^{13,14}. DNA helicases destabilize hydrogen bonds between complementary base pairs of double-stranded DNA in a reaction that is coupled to hydrolysis of nucleoside 5'-triphosphates¹², and these enzymes can translocate along DNA at a rate of several hundred base pairs per second to unwind duplex DNA in a processive manner. Although the ATPase subunits of chromatin-remodelling factors have not been found to possess DNA helicase activity, it is possible that helicases and chromatin-remodelling factors share similarities in their mechanisms of action. For instance, the DNA helicases *Escherichia coli* RecBCD and SV40 large T antigen have been observed to disrupt chromatin structure *in vitro*^{15,16}. Conversely, chromatin-remodelling factors may exhibit properties, such as processive ATP-driven translocation along DNA, that are characteristic of DNA helicases.

In this study, we explore the dynamics of chromatin assembly by ACF. The assembly of nucleosomes can be carried out *in vitro* with purified recombinant ACF, purified recombinant nuclear assembly protein-1 (NAP-1, a core histone chaperone), purified core histones, plasmid DNA and ATP¹¹. This reaction represents the central nucleosome assembly process, which is also linked to other events such as DNA replication and repair^{5,17–19}. Because of the relation of

the ISWI subunit of ACF to DNA helicases, we tested whether ACF is a DNA-tracking enzyme. To this end, we performed template commitment studies on ACF during chromatin assembly. If ACF assembles nucleosomes by tracking processively along DNA, then it would be expected to exhibit commitment to a DNA template during chromatin assembly. These experiments were performed as depicted in the diagram at the bottom of Fig. 1a. First, chromatin assembly was initiated on Template 1 (relaxed circular DNA along with ACF, NAP-1, histones and ATP). Then, 2 min later, Template 2 (relaxed circular DNA along with NAP-1, histones and ATP) was added, and the reaction was allowed to proceed for an additional 0, 20 or 40 min before termination and analysis of DNA supercoiling. These studies involved the use of equivalent masses of two circular DNA templates of 8,117 base pairs (bp) and 3,269 bp ('8K' and '3K', respectively), which can be distinguished in DNA supercoiling assays as Templates 1 and/or 2. The DNA supercoiling gel is shown in Fig. 1a, and the results are summarized in Fig. 1b–d. When both 8K and 3K are added together as Template 1, then we observe the same rate of chromatin assembly on each of the templates (Fig. 1b). When 3K is Template 1 and 8K is Template 2, the rate of chromatin assembly is faster on 3K than on 8K (Fig. 1c). Likewise, the reverse is seen when 8K is Template 1 and 3K is Template 2 (Fig. 1d). As a control, chromatin assembly was not seen in the absence of ACF (Fig. 1a, lanes 16–18).

Next, to investigate the extent of template commitment by ACF, we performed template commitment assays at a lower ACF concentration than that used in Fig. 1, to ensure that ACF is a limiting factor in the assembly reaction (Fig. 2). In these experiments, we used templates of 3 kilobases (kb) ('3K') or 23 kb ('23K') as

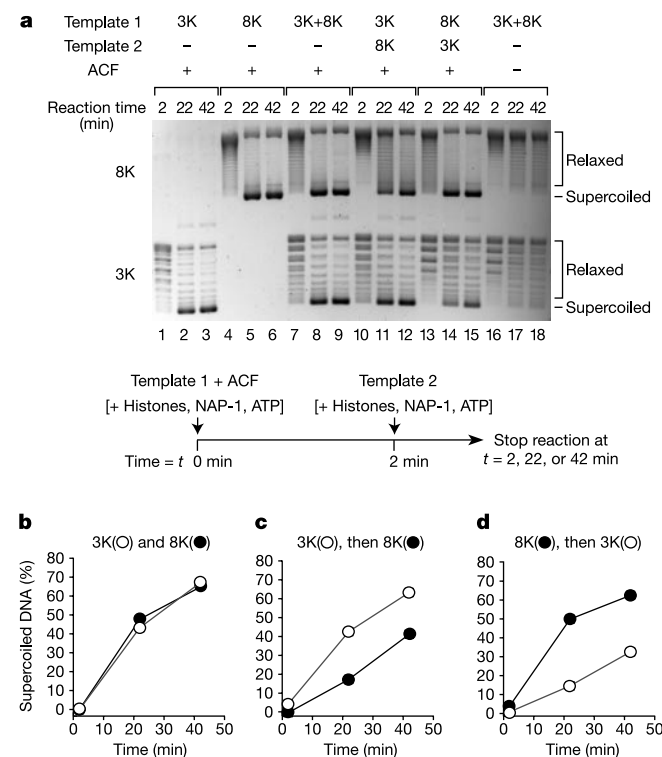


Figure 1 Template commitment by ACF during chromatin assembly. **a**, Assay for template commitment during chromatin assembly. The reaction scheme is illustrated in the diagram. Purified topoisomerase I was also included in the reactions. The extent of nucleosome assembly was monitored by using the DNA supercoiling assay with circular plasmid DNA templates of about 3 kb ('3K') and 8 kb ('8K'). Equivalent masses of 3K and 8K templates were used. **b–d**, Graphs of template commitment results in **a**. The concentration of ACF was the same in each of the reactions, and corresponds to ~1 ACF protomer per 3 kb DNA.

Template 1 and an 8-kb ('8K') Template 2. With either 3K or 23K Template 1, we observed a distinct lag of about 15–30 min in the assembly of 8K Template 2. (As a control, when each of these sets of plasmid DNAs (that is, 3K and 8K, or 23K and 8K) are both added before chromatin assembly, as in Fig. 1b, we observe comparable rates of chromatin assembly from both templates (data not shown). These results reveal that ACF remains committed to the template for an average time of about 15–30 min during chromatin assembly.

We then tested whether template commitment by ACF requires ongoing chromatin assembly (Fig. 3a). We pre-incubated ACF with Template 1 (plus histones and NAP-1) in the absence of ATP for 2 min (in a manner analogous to that in Fig. 1). Hence, all of the components required for chromatin assembly, except for ATP, were present. Next, we added Template 2 (plus histones and NAP-1) and ATP, and allowed the reactions to proceed for 0, 10 or 20 min. These experiments revealed that there was no template commitment by ACF in the absence of ongoing chromatin assembly. Also, ACF template commitment was not observed when ACF + DNA + ATP were combined in the absence of NAP-1 and core histones (data not shown). These findings suggest that template commitment by ACF requires the chromatin assembly machinery to be in an active engaged state and/or the presence of nucleosomes on the template.

We therefore examined whether ACF exhibits a preference for assembly of chromatin onto a template that contains nucleosomes. For these experiments, we used templates, termed 'partial chromatin', that contain an average of 2 nucleosomes per 3-kb template. As depicted in Fig. 3b, we included both 3K 'partial chromatin' and 8K plasmid DNA (both templates were relaxed with topoisomerase I) in chromatin assembly reactions, and did not observe preferential assembly of nucleosomes onto the 'partial chromatin' template. Thus, the ACF-driven chromatin assembly machinery does not seem to exhibit a preference for the assembly of chromatin onto a nucleosome-containing template relative to naked DNA.

We further investigated template commitment by ACF in a single template assay. The extent of nucleosome assembly was monitored by DNA supercoiling analysis, in which individual topoisomers were resolved by two-dimensional agarose gel electrophoresis. We assembled chromatin with a substoichiometric amount of ACF (relative to plasmid DNA molecules) in the presence of topoisomerase

I, and terminated the reactions at an early stage. In this manner, we were able to analyse the initial reaction products under conditions where only a fraction of the plasmids could be associated with ACF. We observed two populations of reaction products (Fig. 4a): one peak corresponded to partially assembled chromatin with an average of about seven negative supercoils, whereas the other peak corresponded to unassembled naked DNA with an average of zero supercoils. These findings indicate that ACF mediates the assembly of at least seven nucleosomes before dissociation from the template, and provide further evidence for template commitment by ACF. We also examined partial chromatin that was reconstituted by salt-gradient dialysis and relaxed by topoisomerase I, and, in contrast, we found a single population of partially reconstituted chromatin species (Fig. 4b).

We also analysed the distribution of nucleosomes on the partially assembled chromatin by using the micrococcal nuclease digestion assay. Chromatin that was partially assembled by ACF (with an average of about 4.6 nucleosomes per template) contained localized arrays of several nucleosomes (Fig. 4a). In contrast, chromatin that was partially reconstituted by salt dialysis (with an average of about 3.8 nucleosomes per template) yielded only mononucleosomes (Fig. 4b), which indicates that the nucleosomes reconstituted by salt dialysis were distributed randomly throughout the template. These results, combined with the ACF template commitment data, suggest that ACF tracks along the DNA and assembles nucleosomes processively.

ACF consists of two subunits: Acf1 and the ISWI ATPase. A single amino-acid substitution in the nucleotide-binding motif of ISWI results in a roughly 20-fold reduction in the ATPase and chromatin assembly activities of the ACF complex (see Supplementary Infor-

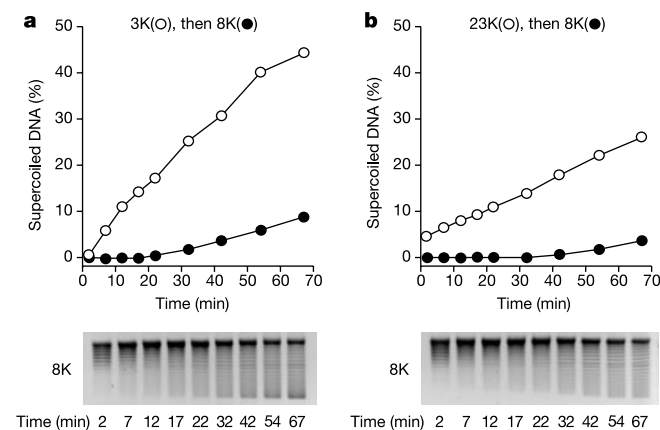


Figure 2 ACF remains committed to the DNA template for about 15–30 min. Template commitment during chromatin assembly was analysed as in Fig. 1. **a**, Template commitment with a 3-kb plasmid ('3K') as Template 1. **b**, Template commitment with a 23-kb ('23K') plasmid as Template 1. When Templates 1 and 2 are added simultaneously, the templates are assembled at comparable rates (data not shown). The extent of commitment to Template 1 (either 3K or 23K) is revealed by the lack of assembly on Template 2 (8K). Equivalent masses of 3K, 8K and 23K templates were used. The concentration of ACF was the same in each reaction, and corresponds to ~0.4 ACF protomers per 3 kb DNA.

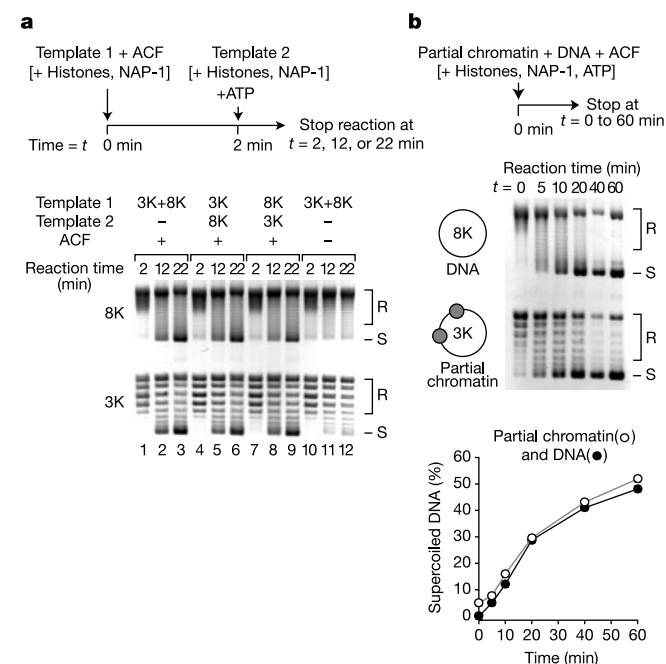


Figure 3 Template commitment by ACF requires continual nucleosome assembly. Chromatin assembly reactions and DNA supercoiling analyses were performed as in Fig. 1. **a**, Pre-incubation of ACF with DNA, core histones and NAP-1 in the absence of chromatin assembly does not result in template commitment. As outlined in the diagram, reactions were performed in a manner analogous to that in Fig. 1. R, relaxed; S, supercoiled. **b**, The presence of nucleosomes on the DNA template does not lead to preferential chromatin assembly by ACF. DNA templates (3 kb) containing an average of two nucleosomes ('partial chromatin') were prepared by salt dialysis methods with purified core histones. These 'partial chromatin' templates were then used in template commitment experiments, as indicated in the diagram. A graph of the results is also shown.

mation). Thus, the hydrolysis of ATP seems to be required for chromatin assembly by ACF. The presence of the ISWI subunit in ACF, the requirement for ATP hydrolysis during chromatin assembly, and the ability of ACF to track along DNA suggested that ACF functions in a manner similar to DNA helicases. We tested the ability of ACF to unwind duplex DNA, but did not observe any detectable strand-separation activity (data not shown). The hydrolysis of ATP by DNA helicases is required for DNA translocation as well as for strand separation²⁰. The *Escherichia coli* Rep and RecBCD DNA helicases have been observed to hydrolyse more than one molecule of ATP per bp of unwound DNA¹². These findings are consistent with a model in which helicases hydrolyse one ATP molecule per bp during translocation. We therefore investigated whether ACF consumes a comparable amount of ATP during chromatin assembly.

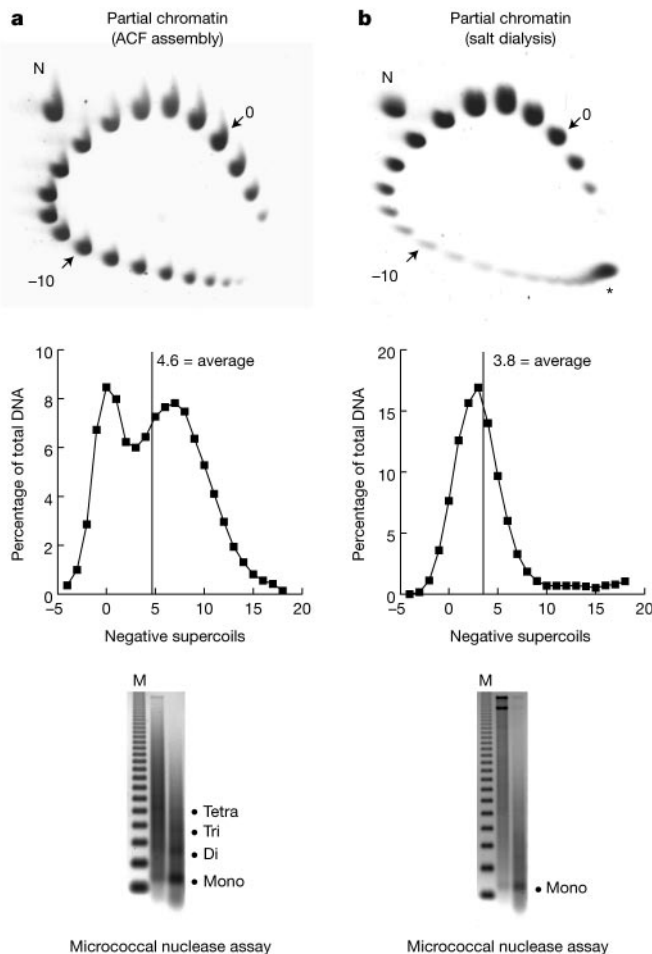


Figure 4 Partial chromatin assembly by ACF reveals template commitment and localized arrays of nucleosomes. **a**, Two-dimensional DNA supercoiling analysis and micrococcal nuclease digestion assay of partial chromatin that is assembled with ACF. Chromatin assembly was performed with purified recombinant dNAP-1 and ACF at ~ 0.2 ACF protomers per 3 kb plasmid DNA. Reactions were terminated at an early stage (10 min reaction time). **b**, Two-dimensional DNA supercoiling analysis and micrococcal nuclease digestion assay of partial chromatin that is reconstituted by salt dialysis at a 5:1 mass ratio of DNA to core histones. In the two-dimensional agarose gels, 0 and -10 indicate the positions of topoisomers with 0 and 10 negative supercoils per template; N specifies the position of nicked DNA; and the asterisk denotes an unknown species of highly negatively supercoiled DNA that fails to be relaxed by excess topoisomerase I and is present in the salt-dialysed chromatin and in the relaxed naked DNA reference (data not shown). The approximate average number of nucleosomes per template is designated by a vertical bar. The positions of DNA derived from mono-, di-, tri-, and tetra-nucleosomes is indicated in the micrococcal nuclease digestion ladders. M is a 123 bp DNA ladder (Gibco BRL).

To this end, we analysed the amount of ATP that is consumed in the assembly reaction. In these experiments, we determined the amount of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ that is hydrolysed under three conditions: (1) ACF-mediated chromatin assembly; (2) incubation of ACF with chromatin (prepared by salt dialysis), which reveals the amount of background ATP hydrolysis by ACF in the presence of chromatin; and (3) incubation of ACF with naked DNA, which indicates the amount of background ATP hydrolysis by ACF in the presence of DNA (Fig. 5a). (In addition, the presence of NAP-1 or topoisomerase I does not affect the ATPase activity of ACF in the presence of DNA or chromatin (data not shown).) In parallel, we determined the total number of nucleosomes in each chromatin sample by quantifying each of the topoisomers in a two-dimensional DNA supercoiling gel (Fig. 5b). In ACF-mediated chromatin assembly reactions, there is a background level of ATP hydrolysis by ACF that is stimulated by chromatin and DNA. To assess the minimal amount of ATP hydrolysis that is expended for the assembly of nucleosomes, we subtracted the amount of ATP consumed in the ACF + chromatin incubation from that consumed in the assembly reaction. Then, division of this value by the total number of nucleosomes in the sample yields the minimal amount of ATP that is hydrolysed specifically for nucleosome assembly, which ranged from 270 ± 50 (at 10 min reaction time) to 380 ± 20 (at 20 min reaction time) ATP molecules per nucleosome assembled. There is a higher efficiency of chromatin assembly (as seen by fewer ATP molecules required for the formation of one nucleosome) at 10 min than at 20 min. It is possible, for instance, that the efficiency of the reaction decreases as the template becomes more fully occupied with nucleosomes (and thus contains less available DNA for nucleosome assembly) at 20 min relative to 10 min reaction time.

It is important to consider that the assembly reaction contains mostly naked DNA at early reaction times. Hence, we also obtained a maximal estimate for the amount of ATP expended for nucleosome assembly by subtraction of the amount of ATP consumed in the ACF + naked DNA incubation from that consumed in the assembly reaction. When this value is divided by the total number of nucleosomes, we obtain the maximal amount of ATP that is hydrolysed for nucleosome assembly, which ranged from 470 ± 70 (at 10 min reaction time) to 640 ± 30 (at 20 min reaction time) ATP molecules per nucleosome assembled. The chromatin has a repeat length of about 170 bp, and so the minimal and maximal values vary from about 1.6 ± 0.3 to 3.8 ± 0.2 molecules of ATP hydrolysed per bp of DNA that is assembled into chromatin.

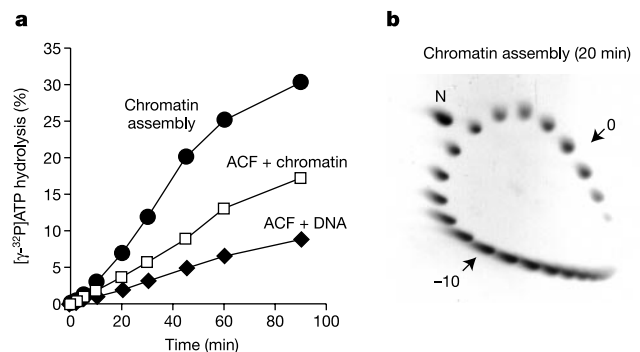


Figure 5 Chromatin assembly by ACF consumes about 270–640 molecules of ATP per nucleosome. **a**, Time course of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ hydrolysis during ACF-mediated chromatin assembly with relaxed plasmid DNA (filled circles), ACF incubation with chromatin (prepared from plasmid DNA and purified histones by salt dialysis techniques; open squares), and ACF incubation with naked DNA (filled diamonds). In these experiments, $n = 4$ and the s.d. of each of the data points did not exceed 0.3 units of the y axis of the plot. We used ~ 0.2 ACF protomers per 3 kb plasmid DNA. **b**, Two-dimensional DNA supercoiling analysis of chromatin assembly at 20 min reaction time. The approximate average number of nucleosomes per template is 8.2.

Thus, the thermodynamic efficiency of chromatin assembly by ACF (2–4 molecules of ATP hydrolysed per bp assembled into chromatin) is similar to that of DNA unwinding by DNA helicases (1–3 molecules of ATP hydrolysed per bp unwound DNA).

The results of this study, along with the presence of the helicase-related ISWI subunit in ACF, implicate a model for ACF-mediated chromatin assembly in which ACF functions as a processive ATP-driven DNA-translocating motor. By analogy to helicases and other DNA-translocating enzymes (for example, polymerases), it is reasonable that the translocation of ACF along the DNA will consume one molecule of ATP per bp DNA. Additional ATP may also be required to disrupt contacts between core histones and NAP-1 as well as to establish proper histone–DNA contacts. In this manner, more than one molecule of ATP per bp might be used during the chromatin assembly process. It is also relevant to consider the potential role of superhelical tension in the DNA during chromatin assembly by ACF. For instance, ISWI and other related ATPases can generate superhelical tension in unconstrained DNA and chromatin²¹. In fact, DNA tracking by ACF may generate transient superhelical tension in DNA. A processive ATP-driven translocation model for nucleosome mobilization has been suggested previously for polymerases and chromatin remodelling factors (see, for example, refs 22–24), and involves the formation of a loop of DNA that is propagated across the surface of the histone octamer. A DNA-tracking mechanism could also be relevant to the nucleosome sliding that has been observed with the chromatin remodelling complexes NURF²⁵, CHRAC^{26,27}, and SWI/SNF²⁸. This study provides evidence that suggests a processive DNA-translocating mechanism for ACF-mediated chromatin assembly. This process involves the use of a considerable amount of ATP. This expenditure of energy provides the cell with the packaging and organization of its genome. □

Methods

Template experiments

The template commitment experiments were performed with three different plasmid DNA templates, as indicated in the figures. Templates 23K, 8K and 3K (pHB23 (~23 kb), pFBI (8,117 bp) and pGIE-0 (3,269 bp), respectively) were prepared by the alkaline lysis method followed by double CsCl gradient centrifugation. The 'partial chromatin' used in Figs 3b and 4b was prepared by salt dialysis¹⁰ from supercoiled pGIE-0 and purified *Drosophila* core histones at 10:1 and 5:1 mass ratios of DNA to core histones, respectively. This procedure resulted in an approximate average number of two or four nucleosomes per 'partial chromatin' molecule. In the two-dimensional DNA supercoiling analysis, electrophoresis in the first direction was performed in the absence of chloroquine. Then, the gel was equilibrated in a buffer containing 1 μ M chloroquine, and electrophoresis was carried out in the second direction. The position of the '0 supercoils' spot was determined by the analysis of fully relaxed naked DNA. The relative intensity of each spot was quantified with the Alpha Imager 2200 digital imaging system (Alpha Innotech).

ATP hydrolysis

In the ATP hydrolysis experiments, chromatin assembly reactions were performed essentially as described¹¹, except that 0.01 μ Ci of [γ -³²P]ATP was added to every reaction, and protein concentrations were increased about fivefold to achieve readily detectable ATP hydrolysis levels. A typical 7- μ l reaction contained 0.04 pmol ACF, 3.7 pmol core histone octamers, 0.20 pmol plasmid DNA (3.2 kb), 30 pmol dNAP-1, 21 nmol ATP, 210 nmol creatine phosphate and 0.1 μ g phosphocreatine kinase. Before the assembly reaction, the plasmid DNA template was relaxed with purified recombinant topoisomerase I. Chromatin templates that were used in the chromatin remodelling reactions were reconstituted by salt dialysis and purified by sucrose gradient sedimentation¹⁰. Both the assembly and remodelling reactions included an excess of topoisomerase I. The hydrolysis of ATP was detected by thin-layer chromatography on cellulose-PEI (SelectoScientific) in TLC buffer (0.8 M acetic acid, 0.8 M LiCl). The amount of total ATP hydrolysed in the reactions was calculated by the equation $([P_i]/[ATP_{tot}]) = -\ln\{1 - ([P_i]/[ATP_{tot}^*])\}$, where $[P_i]$ is the concentration of total inorganic phosphate generated in the reaction (which reflects the amount of ATP hydrolysed), $[ATP_{tot}]$ is the concentration of total ATP in the reaction (which is assumed to remain approximately constant in the presence of the ATP-regenerating system), and $[P_i]/[ATP_{tot}] = [P_i]/([ATP^*] + [P_i])$ is the fraction of [γ -³²P]ATP hydrolysed (shown in Fig. 5a), as determined by quantification of the unhydrolysed [γ -³²P]ATP (ATP*) and [³²P] inorganic phosphate (P_i^*) by using a phosphorimager. We did not observe a transfer of [γ -³²P]phosphate to creatine, and so the ATP regeneration system did not seem to interfere with the determination of the rates of ATP hydrolysis. To confirm the integrity of the chromatin, micrococcal nuclease digestion analysis was performed on the chromatin assembly and chromatin remodelling products after 2 h of reaction time. In all instances tested, the micrococcal nuclease ladders of the

chromatin from both the remodelling and the assembly reactions looked essentially identical (data not shown).

Helicase assays

DNA helicase assays were performed with single-stranded M13 DNA and a complementary oligonucleotide, and helicase activity was observed with SV40 large T antigen (as a positive control; gift of C. Gruss) but not with ACF (data not shown).

Received 30 November 2001; accepted 11 June 2002; doi:10.1038/nature00929.

1. Tsukiyama, T. & Wu, C. Chromatin remodeling and transcription. *Curr. Opin. Genet. Dev.* **7**, 182–191 (1997).
2. Cairns, B. R. Chromatin remodeling machines: similar motors, ulterior motives. *Trends Biochem. Sci.* **23**, 20–25 (1998).
3. Kornberg, R. D. & Lorch, Y. Chromatin-modifying and -remodeling complexes. *Curr. Opin. Genet. Dev.* **9**, 148–151 (1999).
4. Kingston, R. E. & Narlikar, G. J. ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes Dev.* **13**, 2339–2352 (1999).
5. Fyodorov, D. V. & Kadonaga, J. T. The many faces of chromatin remodeling: SWItching beyond transcription. *Cell* **106**, 523–525 (2001).
6. Vignali, M., Hassan, A. H., Neely, K. E. & Workman, J. L. ATP-dependent chromatin-remodeling complexes. *Mol. Cell. Biol.* **20**, 1899–1910 (2000).
7. Fry, C. J. & Peterson, C. L. Chromatin remodeling enzymes: who's on first? *Curr. Biol.* **11**, R185–R197 (2001).
8. Flaus, A. & Owen-Hughes, T. Mechanisms for ATP-dependent chromatin remodeling. *Curr. Opin. Genet. Dev.* **11**, 148–154 (2001).
9. Glikin, G. C., Ruberti, I. & Worcel, A. Chromatin assembly in *Xenopus* oocytes: *in vitro* studies. *Cell* **37**, 33–41 (1984).
10. Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R. & Kadonaga, J. T. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145–155 (1997).
11. Ito, T. *et al.* ACF consists of two subunits, Acl1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev.* **13**, 1529–1539 (1999).
12. Lohman, T. M. & Bjornson, K. P. Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169–214 (1996).
13. Gorbalenya, A. W. & Koonin, E. V. Helicases: amino acid sequence comparisons and structure–function relationships. *Curr. Opin. Struct. Biol.* **3**, 419–429 (1993).
14. Eisen, J. A., Sweder, K. S. & Hanawalt, P. C. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* **23**, 2715–2723 (1995).
15. Eggleston, A. K., O'Neill, T. E., Bradbury, E. M. & Kowalczykowski, S. C. Unwinding of nucleosomal DNA by a DNA helicase. *J. Biol. Chem.* **270**, 2024–2031 (1995).
16. Ramsperger, U. & Stahl, H. Unwinding of chromatin by the SV40 large T antigen DNA helicase. *EMBO J.* **14**, 3215–3225 (1995).
17. Adams, C. R. & Kamakaka, R. T. Chromatin assembly: biochemical identities and genetic redundancy. *Curr. Opin. Genet. Dev.* **9**, 185–190 (1999).
18. Verreault, A. De novo nucleosome assembly: new pieces in an old puzzle. *Genes Dev.* **14**, 1430–1438 (2000).
19. Mello, J. A. & Almouzni, G. The ins and outs of nucleosome assembly. *Curr. Opin. Genet. Dev.* **11**, 136–141 (2001).
20. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999).
21. Havas, K. *et al.* Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell* **103**, 1133–1142 (2000).
22. Studitsky, V. M., Clark, D. J. & Felsenfeld, G. Overcoming a nucleosomal barrier to transcription. *Cell* **83**, 19–27 (1995).
23. Kornberg, R. D. & Lorch, Y. Interplay between chromatin structure and transcription. *Curr. Opin. Cell Biol.* **7**, 371–375 (1995).
24. Pazin, M. J. & Kadonaga, J. T. SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein–DNA interactions? *Cell* **88**, 737–740 (1997).
25. Hamiche, A., Sandaltzopoulos, R., Gdula, D. A. & Wu, C. ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell* **97**, 833–842 (1999).
26. Längst, G., Bonte, E. J., Corona, D. F. V. & Becker, P. B. Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell* **97**, 843–852 (1999).
27. Eberharter, A. *et al.* Acl1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO J.* **20**, 3781–3788 (2001).
28. Whitehouse, I. *et al.* Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**, 784–787 (1999).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements

We thank P. Geiduschek, R. Dutnall, L. Pillus, T. Juven-Gershon, V. Alexiadis, T. Boulay, B. Santos, A. Lusser, J. Huang Parsons and M. Levenstein for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health to J.T.K. D.V.F. is an American Cancer Society Postdoctoral Fellow.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.T.K. (e-mail: jkadonaga@ucsd.edu).

A new line in tubes

Chromatography products — and some gripping news about tubing.

Galaxie

Varian www.varianinc.com

Network software for the laboratory

The Galaxie software program can be used to integrate liquid and gas chromatography instruments from multiple manufacturers into a single computer network, eliminating the need for a dedicated control workstation with each instrument in a laboratory. The system can be expanded to handle hundreds of instruments on one network and is compatible with audit, reporting and security regulations, including US FDA CFR Part 11. Data acquisition and remote instrument control can be performed from any point on the network. Varian estimates that, with the Galaxie network, two or three operators can control 20–30 instruments.

HPS autoclaves

Integra Biosciences

www.integra-biosciences.com

Put a lid on it

Integra's High Pathogen Safe autoclaves offer a sterilization temperature of 140 °C, corresponding to a working pressure of 4.7 bar. They are designed to completely enclose the product, exhaust air and condensate within the unit until sterilization is complete, eliminating the potential for pathogenic material to be released. All enclosed materials can be disposed of as normal waste, without additional processing.



Shut up: HPS autoclaves keep pathogens inside.



A site for sore eyes, from Bel-Art.

Eye wash stations

Bel-Art Products www.belart.com

Emergency solutions

Eye wash stations from Bel-Art can be wall-mounted and are available as single or double stations. They are supplied with 32-ounce bottles filled with sterile, isotonic, buffered eye wash solution, ready for immediate use. The solution rinses away chemical contaminants and other harmful irritants, and will not damage eye or skin tissue with prolonged flushing. Refill bottles are available.

Virtual Column 2

Dionex

www.dionex.com

Free CD-ROM

Virtual Column 2 uses previously acquired retention data to simulate all possible chromatograms in a defined search area. The software uses advanced retention prediction models and embedded data. It can be used in practical method development, education and training, refinement of existing methods and exploratory investigations. The software CD-ROM is available free of charge from Dionex.

AM200 leaf meter

ADC BioScientific

www.adc.co.uk

A worm's eye view

The AM200 is a lightweight, portable hand-scanner that can perform up to 3,000 non-destructive leaf area measurements between recharging of the batteries. Measurements can be made on the integral scan board or on an independent scan surface. It is suitable for measurement of intact long and wide leaves, and, with a resolution down to 0.065 mm², can also measure very small leaves. Up to 2,000 sets of measure-

ments and leaf images can be stored in its memory. The device can also be used to measure areas of disease or discoloration while leaves are still connected to the plant.

Process Separations website

Bio-Rad www.bio-rad.com

For all things chromatography

Bio-Rad has launched a website dedicated to process chromatography. Users can gain access to process chromatography experts, exchange information with peers via a bulletin board, download product bulletins, manuals and application notes and request quotes. The site operates in English and French, and visitors are not required to register. It can be found at www.bio-rad.com/process/.

Zetalif

ESA Analytical

www.esainc.com

Exciting new range

Zetalif laser-induced fluorescence detectors are compatible with existing high-performance liquid chromatography and capillary electrophoresis systems and include a broad range of lasers, from ultraviolet to infrared, as well as systems with deep ultraviolet lasers. The detector with 266 nm excitation can be used to monitor the native fluorescence of



AM200: turn over a new measuring device.

a variety of molecules at nanomolar to picomolar concentrations without the need for derivatization. The 266 nm system can also be used for the detection of common amino acid derivatives.

FCPC Chromatograph

Kromaton Technologies www.kromaton.com

No support for this one

Kromaton's fast centrifugal partition chromatography (FCPC) is a mono-axial apparatus with two high-pressure rotary seals. It has no solid support, but relies on two immiscible liquid phases, one of which is held stationary in the column by centrifugal force. This eliminates loss of compounds as a result of irreversible adsorption by a solid support and denaturation of fragile biological compounds. By selecting either the stationary or the mobile phase, researchers can isolate compounds of widely varying structure and polarity with the same column. It can be used to extract, fractionate and purify a wide variety of organic and inorganic compounds.

Grippers

Omnifit www.omnifit.com

Well-connected

Omnifit's grippers for connecting laboratory tubing offer zero dead volume at pressures up to 1,000 psi. They can be hand-tightened without the need for a special tool and will not compress or twist the tubing if over-tightened. The grippers have a chemically inert washer press-fitted in a stainless steel housing, providing a leak-free connection with minimum disturbance to the fluid path. They can be connected and disconnected repeatedly with no deterioration to the sealing face.



Get a grip: Omnifit's approach to making tube connections.



Two into one: Fast Pak microarrays.

Fast Pak 2 protein microarray kit

Schleicher & Schuell Bioscience

www.s-und-s.de

Twice as nice

The Fast Pak kit from Schleicher & Schuell Bioscience enables researchers to investigate two different microarrays or samples on a single slide. The slides in the kit have two separate 20 × 20 mm pads and corresponding incubation chambers, allowing parallel analysis. Protein microarrays are stable for more than 22 months using the Fast Pak array buffer. The kit can be used with existing arrayers and scanners and is compatible with fluorescent, chemiluminescent, colorimetric and radioactive detection.

SE-200 tubing

Tygon www.tygon.com

Going around the bend, clearly

Tygon's SE-200 tubing has an inert FEP liner that can withstand aggressive chemicals and minimizes the potential for fluid contamination during transfers. The jacket is made of clear, flexible Tygon that can accommodate tight bends. Applications include pharmaceutical processing, ink and toner lines, beverage dispensing and high-purity fluid and corrosive chemical transfer.

Lab Solutions Version 2

Shimadzu Scientific Instruments

www.shimadzu.com

No more macros

This software release from Shimadzu is designed to increase productivity and facilitate compliance with FDA requirements for CFR2 Part 11. It features 'one-window' access to simplify



Tube travel: the latest line from Tygon.

operation and put total control of the system on one screen. Reports can be customized with no macros required and reporting functions comply with USEPA, ASTM and NIDA requirements for format and completeness. The software supports all of the sensitivity and speed functions of the GCMS-QP2010, with SIM and scan modes possible in a single analytical run. A new batch table design processes up to 1,000 samples per batch, with 1,000 batches in the sample queue.

Atlas 2002-08-04

Thermo LabSystems

www.thermolabsystems.com

Software for the lab controller

Atlas is a scaleable chromatography data system that can be used on a single workstation as well as in networked laboratories with multiple users and channels. The 2002 version adds a series of improvements, including control of the Agilent 5890 and 6890, enhanced control of the Agilent 1100, including status monitoring, control of the Thermo Finnigan SpectraSystem modules P4000, AS3000/AS3500 and UV2000 and of A2D Data Server instruments. Other features include customizable toolbars and instrument icons, calibration and service status reports, user credential authentication and the capacity to add a specific audit when modifying a run.

These notes are compiled in the Nature office from information provided by the manufacturers. Press releases to be considered for publication can be sent to newproducts@nature.com.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Silje Opstrup (4994)

France/ Belgium:

Amelie Pequignot (4974)

Production Manager: Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Switzerland:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikatamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

e-mail: k.johnson@naturejpn.com

Asia-Pacific Advertising Manager:

Keyn Johnson

naturejobs

In search of the truth

In an ideal world, lab heads would adopt a 'truth in advertising' policy when it comes to academic career prospects for new graduates. This they would fulfil in a welcoming speech, as well as by posting a prominent message on bulletin boards next to the yellowing *Far Side* cartoons. A potted version, suitable for adopting, adapting or ignoring is as follows:

"Welcome to my lab. I managed to rise to this position through a combination of talent, diligence and luck — the last two of which were the most important. I'm sorry to inform you that the odds of landing a similar position are not great. Although almost all of you probably want to become tenured professors, less than half of you are likely to make it. There just aren't that many open positions, and indications are that there never will be. Sure, your government has been talking up increased funding for the sciences, but much of that money — if it ever really materializes — is likely to go on infrastructure (governments love to build things) or on bigger grants for people already in the business.

"Also, if I haven't yet scared you off the tenure track, please let me take this opportunity to try. It's a long, long, road. I, fortunately, only had to do one short postdoc. You, on the other hand, may need to do several longer ones — unless you pick a good principal investigator and get lucky and publish a stellar paper your first time out. Sure, it can happen, but please don't count on it.

"But fear not. Even if you don't get a tenure-track position, you have alternatives. Industry is not necessarily an evil entity — I've heard that some companies actually let you publish. Also, more research institutions are moving towards short-term contracts; this is not necessarily a bad thing. And finally, some of you may find success in non-traditional areas such as law, finance and the media. If you land one of these positions, you may be entitled to feel as smug as I."

Paul Smaglik

Naturejobs editor



Contents

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS